NAME

perluniprops - Index of Unicode Version 13.0.0 character properties in Perl

DESCRIPTION

This document provides information about the portion of the Unicode database that deals with character properties, that is the portion that is defined on single code points. ("Other information in the Unicode data base" below briefly mentions other data that Unicode provides.)

Perl can provide access to all non-provisional Unicode character properties, though not all are enabled by default. The omitted ones are the Unihan properties (accessible via the CPAN module Unicode::Unihan) and certain deprecated or Unicode-internal properties. (An installation may choose to recompile Perl's tables to change this. See "Unicode character properties that are NOT accepted by Perl".)

For most purposes, access to Unicode properties from the Perl core is through regular expression matches, as described in the next section. For some special purposes, and to access the properties that are not suitable for regular expression matching, all the Unicode character properties that Perl handles are accessible via the standard Unicode::UCD module, as described in the section "Properties accessible through Unicode::UCD".

Perl also provides some additional extensions and short-cut synonyms for Unicode properties.

This document merely lists all available properties and does not attempt to explain what each property really means. There is a brief description of each Perl extension; see "Other Properties" in perlunicode for more information on these. There is some detail about Blocks, Scripts, General_Category, and Bidi_Class in perlunicode, but to find out about the intricacies of the official Unicode properties, refer to the Unicode standard. A good starting place is <http://www.unicode.org/reports/tr44/>.

Note that you can define your own properties; see "User-Defined Character Properties" in perlunicode.

Properties accessible through "\p{}" and "\P{}"

The Perl regular expression "\p{}" and "\P{}" constructs give access to most of the Unicode character properties. The table below shows all these constructs, both single and compound forms.

Compound forms consist of two components, separated by an equals sign or a colon. The first component is the property name, and the second component is the particular value of the property to match against, for example, "\p{Script_Extensions: Greek}" and "\p{Script_Extensions=Greek}" both mean to match characters whose Script_Extensions property value is Greek. ("Script_Extensions" is an improved version of the "Script" property.)

Single forms, like "\p{Greek}", are mostly Perl-defined shortcuts for their equivalent compound forms. The table shows these equivalences. (In our example, "\p{Greek}" is a just a shortcut for "\p{Script_Extensions=Greek}"). There are also a few Perl-defined single forms that are not shortcuts for a compound form. One such is "\p{Word}". These are also listed in the table.

In parsing these constructs, Perl always ignores Upper/lower case differences everywhere within the {braces}. Thus "\p{Greek}" means the same thing as "\p{greek}". But note that changing the case of the "p" or "P" before the left brace completely changes the meaning of the construct, from "match" (for "\p{}") to "doesn't match" (for "\P{}"). Casing in this document is for improved legibility.

Also, white space, hyphens, and underscores are normally ignored everywhere between the {braces}, and hence can be freely added or removed even if the "/x" modifier hasn't been specified on the regular expression. But in the table below a 'T' at the beginning of an entry means that tighter (stricter) rules are used for that entry:

Single form ("\p{name}") tighter rules:
    White space, hyphens, and underscores ARE significant except for:

    ⊕   white space adjacent to a non-word character

    ⊕   underscores separating digits in numbers

    That means, for example, that you can freely add or remove white space adjacent to (but within) the braces without affecting the meaning.

Compound form ("\p{name=value}" or "\p{name:value}") tighter rules:
    The tighter rules given above for the single form apply to everything to the right of the colon or equals; the looser rules still apply to everything to the left.

    That means, for example, that you can freely add or remove white space adjacent to (but within) the braces and the colon or equal sign.

Some properties are considered obsolete by Unicode, but still available. There are several varieties of obsolescence:

Stabilized
    A property may be stabilized. Such a determination does not indicate that the property should or should not be used; instead it is a declaration that the property will not be maintained nor extended for newly encoded characters. Such properties are marked with an 'S' in the table.

Deprecated

> A property may be deprecated, perhaps because its original intent has been replaced by another property, or because its specification was somehow defective. This means that its use is strongly discouraged, so much so that a warning will be issued if used, unless the regular expression is in the scope of a "no warnings 'deprecated'" statement. A 'D' flags each such entry in the table, and the entry there for the longest, most descriptive version of the property will give the reason it is deprecated, and perhaps advice. Perl may issue such a warning, even for properties that aren't officially deprecated by Unicode, when there used to be characters or code points that were matched by them, but no longer. This is to warn you that your program may not work like it did on earlier Unicode releases.
>
> A deprecated property may be made unavailable in a future Perl version, so it is best to move away from them.
>
> A deprecated property may also be stabilized, but this fact is not shown.

Obsolete

> Properties marked with an 'O' in the table are considered (plain) obsolete. Generally this designation is given to properties that Unicode once used for internal purposes (but not any longer).

Discouraged

> This is not actually a Unicode-specified obsolescence, but applies to certain Perl extensions that are present for backwards compatibility, but are discouraged from being used. These are not obsolete, but their meanings are not stable. Future Unicode versions could force any of these extensions to be removed without warning, replaced by another property with the same name that means something different. An 'X' flags each such entry in the table. Use the equivalent shown instead.
>
> In particular, matches in the Block property have single forms defined by Perl that begin with "In_", ""Is_", or even with no prefix at all, Like all **DISCOURAGED** forms, these are not stable. For example, "\p{Block=Deseret}" can currently be written as "\p{In_Deseret}", "\p{Is_Deseret}", or "\p{Deseret}". But, a new Unicode version may come along that would force Perl to change the meaning of one or more of these, and your program would no longer be correct. Currently there are no such conflicts with the form that begins "In_", but there are many with the other two shortcuts, and Unicode continues to define new properties that begin with "In", so it's quite possible that a conflict will occur in the future. The compound form is guaranteed to not become obsolete, and its meaning is clearer anyway. See "Blocks" in perlunicode for more information about this.

User-defined properties must begin with "In" or "Is". These override any Unicode property of the same name.

The table below has two columns. The left column contains the "\p{}" constructs to look up, possibly preceded by the flags mentioned above; and the right column contains information about them, like a description, or synonyms. The table shows both the single and compound forms for each property that has them. If the left column is a short name for a property, the right column will give its longer, more descriptive name; and if the left column is the longest name, the right column will show any equivalent shortest name, in both single and compound forms if applicable.

If braces are not needed to specify a property (e.g., "\pL"), the left column contains both forms, with and without braces.

The right column will also caution you if a property means something different than what might normally be expected.

All single forms are Perl extensions; a few compound forms are as well, and are noted as such.

Numbers in (parentheses) indicate the total number of Unicode code points matched by the property. For the entries that give the longest, most descriptive version of the property, the count is followed by a list of some of the code points matched by it. The list includes all the matched characters in the 0-255 range, enclosed in the familiar [brackets] the same as a regular expression bracketed character class. Following that, the next few higher matching ranges are also given. To avoid visual ambiguity, the SPACE character is represented as "\x20".

For emphasis, those properties that match no code points at all are listed as well in a separate section following the table.

Most properties match the same code points regardless of whether "/i" case-insensitive matching is specified or not. But a few properties are affected. These are shown with the notation "(/i= *other_property*)" in the second column. Under case-insensitive matching they match the same code pode points as the property *other_property*.

There is no description given for most non-Perl defined properties (See <http://www.unicode.org/reports/tr44/> for that).

For compactness, '*****' is used as a wildcard instead of showing all possible combinations. For example, entries like:

  \p{Gc: *}   \p{General_Category: *}

mean that 'Gc' is a synonym for 'General_Category', and anything that is valid for the latter is also valid for the former. Similarly,

 \p{Is_*}   \p{*}

means that if and only if, for example, "\p{Foo}" exists, then "\p{Is_Foo}" and "\p{IsFoo}" are also valid and all mean the same thing. And similarly, "\p{Foo=Bar}" means the same as "\p{Is_Foo=Bar}" and "\p{IsFoo=Bar}". "*" here is restricted to something not beginning with an underscore.

Also, in binary properties, 'Yes', 'T', and 'True' are all synonyms for 'Y'. And 'No', 'F', and 'False' are all synonyms for 'N'. The table shows 'Y*' and 'N*' to indicate this, and doesn't have separate entries for the other possibilities. Note that not all properties which have values 'Yes' and 'No' are binary, and they have all their values spelled out without using this wild card, and a "NOT" clause in their description that highlights their not being binary. These also require the compound form to match them, whereas true binary properties have both single and compound forms available.

Note that all non-essential underscores are removed in the display of the short names below.

Legend summary:

* is a wild-card
(\d+) in the info column gives the number of Unicode code points matched by this property.
D means this is deprecated.
O means this is obsolete.
S means this is stabilized.
T means tighter (stricter) name matching applies.
X means use of this form is discouraged, and may not be stable.

```
    NAME                    INFO


 \p{Adlam}             \p{Script_Extensions=Adlam} (Short:
                 \p{Adlm}; NOT \p{Block=Adlam}) (89)
 \p{Adlm}              \p{Adlam} (= \p{Script_Extensions=Adlam})
                 (NOT \p{Block=Adlam}) (89)
X \p{Aegean_Numbers}     \p{Block=Aegean_Numbers} (64)
T \p{Age: 1.1}         \p{Age=V1_1} (33_979)
 \p{Age: V1_1}          Code point's usage introduced in version
                 1.1 (33_979: U+0000..01F5, U+01FA..0217,
                 U+0250..02A8, U+02B0..02DE,
```

                    U+02E0..02E9, U+0300..0345 ...)
    T \p{Age: 2.0}   \p{Age=V2_0} (144_521)
     \p{Age: V2_0}   Code point's usage was introduced in
                    version 2.0; See also Property
                    'Present_In' (144_521: U+0591..05A1,
                    U+05A3..05AF, U+05C4, U+0F00..0F47,
                    U+0F49..0F69, U+0F71..0F8B ...)
    T \p{Age: 2.1}   \p{Age=V2_1} (2)
     \p{Age: V2_1}   Code point's usage was introduced in
                    version 2.1; See also Property
                    'Present_In' (2: U+20AC, U+FFFC)
    T \p{Age: 3.0}   \p{Age=V3_0} (10_307)
     \p{Age: V3_0}   Code point's usage was introduced in
                    version 3.0; See also Property
                    'Present_In' (10_307: U+01F6..01F9,
                    U+0218..021F, U+0222..0233,
                    U+02A9..02AD, U+02DF, U+02EA..02EE ...)
    T \p{Age: 3.1}   \p{Age=V3_1} (44_978)
     \p{Age: V3_1}   Code point's usage was introduced in
                    version 3.1; See also Property
                    'Present_In' (44_978: U+03F4..03F5,
                    U+FDD0..FDEF, U+10300..1031E,
                    U+10320..10323, U+10330..1034A,
                    U+10400..10425 ...)
    T \p{Age: 3.2}   \p{Age=V3_2} (1016)
     \p{Age: V3_2}   Code point's usage was introduced in
                    version 3.2; See also Property
                    'Present_In' (1016: U+0220, U+034F,
                    U+0363..036F, U+03D8..03D9, U+03F6,
                    U+048A..048B ...)
    T \p{Age: 4.0}   \p{Age=V4_0} (1226)
     \p{Age: V4_0}   Code point's usage was introduced in
                    version 4.0; See also Property
                    'Present_In' (1226: U+0221,
                    U+0234..0236, U+02AE..02AF,
                    U+02EF..02FF, U+0350..0357, U+035D..035F
                    ...)
    T \p{Age: 4.1}   \p{Age=V4_1} (1273)
     \p{Age: V4_1}   Code point's usage was introduced in
                    version 4.1; See also Property

                'Present_In' (1273: U+0237..0241,
                U+0358..035C, U+03FC..03FF,
                U+04F6..04F7, U+05A2, U+05C5..05C7 ...)

T \p{Age: 5.0}   \p{Age=V5_0} (1369)
 \p{Age: V5_0}   Code point's usage was introduced in
                version 5.0; See also Property
                'Present_In' (1369: U+0242..024F,
                U+037B..037D, U+04CF, U+04FA..04FF,
                U+0510..0513, U+05BA ...)

T \p{Age: 5.1}   \p{Age=V5_1} (1624)
 \p{Age: V5_1}   Code point's usage was introduced in
                version 5.1; See also Property
                'Present_In' (1624: U+0370..0373,
                U+0376..0377, U+03CF, U+0487,
                U+0514..0523, U+0606..060A ...)

T \p{Age: 5.2}   \p{Age=V5_2} (6648)
 \p{Age: V5_2}   Code point's usage was introduced in
                version 5.2; See also Property
                'Present_In' (6648: U+0524..0525,
                U+0800..082D, U+0830..083E, U+0900,
                U+094E, U+0955 ...)

T \p{Age: 6.0}   \p{Age=V6_0} (2088)
 \p{Age: V6_0}   Code point's usage was introduced in
                version 6.0; See also Property
                'Present_In' (2088: U+0526..0527,
                U+0620, U+065F, U+0840..085B, U+085E,
                U+093A..093B ...)

T \p{Age: 6.1}   \p{Age=V6_1} (732)
 \p{Age: V6_1}   Code point's usage was introduced in
                version 6.1; See also Property
                'Present_In' (732: U+058F, U+0604,
                U+08A0, U+08A2..08AC, U+08E4..08FE,
                U+0AF0 ...)

T \p{Age: 6.2}   \p{Age=V6_2} (1)
 \p{Age: V6_2}   Code point's usage was introduced in
                version 6.2; See also Property
                'Present_In' (1: U+20BA)

T \p{Age: 6.3}   \p{Age=V6_3} (5)
 \p{Age: V6_3}   Code point's usage was introduced in
                version 6.3; See also Property

　　　　　　　　　　　　　'Present_In' (5: U+061C, U+2066..2069)

T \p{Age: 7.0}　　　　\p{Age=V7_0} (2834)
 \p{Age: V7_0}　　　　Code point's usage was introduced in
　　　　　　　　　　version 7.0; See also Property
　　　　　　　　　　'Present_In' (2834: U+037F,
　　　　　　　　　　U+0528..052F, U+058D..058E, U+0605,
　　　　　　　　　　U+08A1, U+08AD..08B2 ...)

T \p{Age: 8.0}　　　　\p{Age=V8_0} (7716)
 \p{Age: V8_0}　　　　Code point's usage was introduced in
　　　　　　　　　　version 8.0; See also Property
　　　　　　　　　　'Present_In' (7716: U+08B3..08B4,
　　　　　　　　　　U+08E3, U+0AF9, U+0C5A, U+0D5F, U+13F5
　　　　　　　　　　...)

T \p{Age: 9.0}　　　　\p{Age=V9_0} (7500)
 \p{Age: V9_0}　　　　Code point's usage was introduced in
　　　　　　　　　　version 9.0; See also Property
　　　　　　　　　　'Present_In' (7500: U+08B6..08BD,
　　　　　　　　　　U+08D4..08E2, U+0C80, U+0D4F,
　　　　　　　　　　U+0D54..0D56, U+0D58..0D5E ...)

T \p{Age: 10.0}　　　　\p{Age=V10_0} (8518)
 \p{Age: V10_0}　　　　Code point's usage was introduced in
　　　　　　　　　　version 10.0; See also Property
　　　　　　　　　　'Present_In' (8518: U+0860..086A,
　　　　　　　　　　U+09FC..09FD, U+0AFA..0AFF, U+0D00,
　　　　　　　　　　U+0D3B..0D3C, U+1CF7 ...)

T \p{Age: 11.0}　　　　\p{Age=V11_0} (684)
 \p{Age: V11_0}　　　　Code point's usage was introduced in
　　　　　　　　　　version 11.0; See also Property
　　　　　　　　　　'Present_In' (684: U+0560, U+0588,
　　　　　　　　　　U+05EF, U+07FD..07FF, U+08D3, U+09FE ...)

T \p{Age: 12.0}　　　　\p{Age=V12_0} (554)
 \p{Age: V12_0}　　　　Code point's usage was introduced in
　　　　　　　　　　version 12.0; See also Property
　　　　　　　　　　'Present_In' (554: U+0C77, U+0E86,
　　　　　　　　　　U+0E89, U+0E8C, U+0E8E..0E93, U+0E98 ...)

T \p{Age: 12.1}　　　　\p{Age=V12_1} (1)
 \p{Age: V12_1}　　　　Code point's usage was introduced in
　　　　　　　　　　version 12.1; See also Property
　　　　　　　　　　'Present_In' (1: U+32FF)

T \p{Age: 13.0}　　　　\p{Age=V13_0} (5930)

\p{Age: V13_0}   Code point's usage was introduced in
                version 13.0; See also Property
                'Present_In' (5930: U+08BE..08C7,
                U+0B55, U+0D04, U+0D81, U+1ABF..1AC0,
                U+2B97 ...)
\p{Age: NA}   \p{Age=Unassigned} (830_606 plus all
                above-Unicode code points)
\p{Age: Unassigned}   Code point's usage has not been assigned
                in any Unicode release thus far.
(Short: \p{Age=NA}) (830_606 plus all above-Unicode code points:
                U+0378..0379, U+0380..0383, U+038B,
                U+038D, U+03A2, U+0530 ...)
\p{Aghb}   \p{Caucasian_Albanian} (=
                \p{Script_Extensions=
                Caucasian_Albanian}) (NOT \p{Block=
                Caucasian_Albanian}) (53)
\p{AHex}   \p{PosixXDigit} (= \p{ASCII_Hex_Digit=Y})
                (22)
\p{AHex: *}   \p{ASCII_Hex_Digit: *}
\p{Ahom}   \p{Script_Extensions=Ahom} (NOT \p{Block=
                Ahom}) (58)
X \p{Alchemical}   \p{Alchemical_Symbols} (= \p{Block=
                Alchemical_Symbols}) (128)
X \p{Alchemical_Symbols} \p{Block=Alchemical_Symbols} (Short:
                \p{InAlchemical}) (128)
\p{All}   All code points, including those above
                Unicode. Same as qr/./s (1_114_112 plus
                all above-Unicode code points:
                U+0000..infinity)
\p{Alnum}   \p{XPosixAlnum} (133_525)
\p{Alpha}   \p{XPosixAlpha} (= \p{Alphabetic=Y})
                (132_875)
\p{Alpha: *}   \p{Alphabetic: *}
\p{Alphabetic}   \p{XPosixAlpha} (= \p{Alphabetic=Y})
                (132_875)
\p{Alphabetic: N*}   (Short: \p{Alpha=N}, \P{Alpha}) (981_237
                plus all above-Unicode code points:
                [\x00-\x20!\"#\$\%&\'\(\)*+,\-.\/0-9:;<=
                >?\@\[\\\]\^_'\{\|\}~\x7f-\xa9\xab-\xb4
                \xb6-\xb9\xbb-\xbf\xd7\xf7],

```
                       U+02C2..02C5, U+02D2..02DF,
                       U+02E5..02EB, U+02ED, U+02EF..0344 ...)
  \p{Alphabetic: Y*}    (Short: \p{Alpha=Y}, \p{Alpha}) (132_875:
                       [A-Za-z\xaa\xb5\xba\xc0-\xd6\xd8-\xf6
                       \xf8-\xff], U+0100..02C1, U+02C6..02D1,
                       U+02E0..02E4, U+02EC, U+02EE ...)
X \p{Alphabetic_PF}     \p{Alphabetic_Presentation_Forms} (=
                       \p{Block=Alphabetic_Presentation_Forms})
                       (80)
X \p{Alphabetic_Presentation_Forms} \p{Block=
                       Alphabetic_Presentation_Forms} (Short:
                       \p{InAlphabeticPF}) (80)
  \p{Anatolian_Hieroglyphs} \p{Script_Extensions=
                       Anatolian_Hieroglyphs} (Short: \p{Hluw};
                       NOT \p{Block=Anatolian_Hieroglyphs})
                       (583)
X \p{Ancient_Greek_Music} \p{Ancient_Greek_Musical_Notation} (=
                       \p{Block=
                       Ancient_Greek_Musical_Notation}) (80)
X \p{Ancient_Greek_Musical_Notation} \p{Block=
                       Ancient_Greek_Musical_Notation} (Short:
                       \p{InAncientGreekMusic}) (80)
X \p{Ancient_Greek_Numbers} \p{Block=Ancient_Greek_Numbers} (80)
X \p{Ancient_Symbols}    \p{Block=Ancient_Symbols} (64)
  \p{Any}               All Unicode code points (1_114_112:
                       U+0000..10FFFF)
  \p{Arab}              \p{Arabic} (= \p{Script_Extensions=
                       Arabic}) (NOT \p{Block=Arabic}) (1335)
  \p{Arabic}            \p{Script_Extensions=Arabic} (Short:
                       \p{Arab}; NOT \p{Block=Arabic}) (1335)
X \p{Arabic_Ext_A}      \p{Arabic_Extended_A} (= \p{Block=
                       Arabic_Extended_A}) (96)
X \p{Arabic_Extended_A}  \p{Block=Arabic_Extended_A} (Short:
                       \p{InArabicExtA}) (96)
X \p{Arabic_Math}       \p{Arabic_Mathematical_Alphabetic_Symbols}
                       (= \p{Block=
                       Arabic_Mathematical_Alphabetic_Symbols})
                       (256)
X \p{Arabic_Mathematical_Alphabetic_Symbols} \p{Block=
                       Arabic_Mathematical_Alphabetic_Symbols}
```

```
                      (Short: \p{InArabicMath}) (256)
   X \p{Arabic_PF_A}        \p{Arabic_Presentation_Forms_A} (=
                      \p{Block=Arabic_Presentation_Forms_A})
                      (688)
   X \p{Arabic_PF_B}        \p{Arabic_Presentation_Forms_B} (=
                      \p{Block=Arabic_Presentation_Forms_B})
                      (144)
   X \p{Arabic_Presentation_Forms_A} \p{Block=
                      Arabic_Presentation_Forms_A} (Short:
                      \p{InArabicPFA}) (688)
   X \p{Arabic_Presentation_Forms_B} \p{Block=
                      Arabic_Presentation_Forms_B} (Short:
                      \p{InArabicPFB}) (144)
   X \p{Arabic_Sup}         \p{Arabic_Supplement} (= \p{Block=
                      Arabic_Supplement}) (48)
   X \p{Arabic_Supplement}   \p{Block=Arabic_Supplement} (Short:
                      \p{InArabicSup}) (48)
    \p{Armenian}           \p{Script_Extensions=Armenian} (Short:
                      \p{Armn}; NOT \p{Block=Armenian}) (96)
    \p{Armi}               \p{Imperial_Aramaic} (=
                      \p{Script_Extensions=Imperial_Aramaic})
                      (NOT \p{Block=Imperial_Aramaic}) (31)
    \p{Armn}               \p{Armenian} (= \p{Script_Extensions=
                      Armenian}) (NOT \p{Block=Armenian}) (96)
   X \p{Arrows}            \p{Block=Arrows} (112)
    \p{ASCII}              \p{Block=Basic_Latin} (128)
    \p{ASCII_Hex_Digit}     \p{PosixXDigit} (= \p{ASCII_Hex_Digit=Y})
                      (22)
    \p{ASCII_Hex_Digit: N*} (Short: \p{AHex=N}, \P{AHex}) (1_114_090
                      plus all above-Unicode code points:
                      [\x00-\x20!\"#\$\%&\'\(\)*+,\-.\/:;<=>?
                      \@G-Z\[\\\]\^_'g-z\{\|\}~\x7f-\xff],
                      U+0100..infinity)
    \p{ASCII_Hex_Digit: Y*} (Short: \p{AHex=Y}, \p{AHex}) (22: [0-9A-
                      Fa-f])
    \p{Assigned}          All assigned code points (283_440:
                      U+0000..0377, U+037A..037F,
                      U+0384..038A, U+038C, U+038E..03A1,
                      U+03A3..052F ...)
    \p{Avestan}            \p{Script_Extensions=Avestan} (Short:
```

```
                    \p{Avst}; NOT \p{Block=Avestan}) (61)
    \p{Avst}          \p{Avestan} (= \p{Script_Extensions=
                    Avestan}) (NOT \p{Block=Avestan}) (61)
    \p{Bali}          \p{Balinese} (= \p{Script_Extensions=
                    Balinese}) (NOT \p{Block=Balinese}) (121)
    \p{Balinese}        \p{Script_Extensions=Balinese} (Short:
                    \p{Bali}; NOT \p{Block=Balinese}) (121)
    \p{Bamu}          \p{Bamum} (= \p{Script_Extensions=Bamum})
                    (NOT \p{Block=Bamum}) (657)
    \p{Bamum}          \p{Script_Extensions=Bamum} (Short:
                    \p{Bamu}; NOT \p{Block=Bamum}) (657)
  X \p{Bamum_Sup}        \p{Bamum_Supplement} (= \p{Block=
                    Bamum_Supplement}) (576)
  X \p{Bamum_Supplement}    \p{Block=Bamum_Supplement} (Short:
                    \p{InBamumSup}) (576)
  X \p{Basic_Latin}        \p{ASCII} (= \p{Block=Basic_Latin}) (128)
    \p{Bass}          \p{Bassa_Vah} (= \p{Script_Extensions=
                    Bassa_Vah}) (NOT \p{Block=Bassa_Vah})
                    (36)
    \p{Bassa_Vah}        \p{Script_Extensions=Bassa_Vah} (Short:
                    \p{Bass}; NOT \p{Block=Bassa_Vah}) (36)
    \p{Batak}          \p{Script_Extensions=Batak} (Short:
                    \p{Batk}; NOT \p{Block=Batak}) (56)
    \p{Batk}          \p{Batak} (= \p{Script_Extensions=Batak})
                    (NOT \p{Block=Batak}) (56)
    \p{Bc: *}          \p{Bidi_Class: *}
    \p{Beng}          \p{Bengali} (= \p{Script_Extensions=
                    Bengali}) (NOT \p{Block=Bengali}) (113)
    \p{Bengali}         \p{Script_Extensions=Bengali} (Short:
                    \p{Beng}; NOT \p{Block=Bengali}) (113)
    \p{Bhaiksuki}        \p{Script_Extensions=Bhaiksuki} (Short:
                    \p{Bhks}; NOT \p{Block=Bhaiksuki}) (97)
    \p{Bhks}          \p{Bhaiksuki} (= \p{Script_Extensions=
                    Bhaiksuki}) (NOT \p{Block=Bhaiksuki})
                    (97)
    \p{Bidi_C}         \p{Bidi_Control} (= \p{Bidi_Control=Y})
                    (12)
    \p{Bidi_C: *}        \p{Bidi_Control: *}
    \p{Bidi_Class: AL}     \p{Bidi_Class=Arabic_Letter} (1698)
    \p{Bidi_Class: AN}     \p{Bidi_Class=Arabic_Number} (61)
```

\p{Bidi_Class: Arabic_Letter} (Short: \p{Bc=AL}) (1698: U+0608,
          U+060B, U+060D, U+061B..064A,
          U+066D..066F, U+0671..06D5 ...)
\p{Bidi_Class: Arabic_Number} (Short: \p{Bc=AN}) (61:
          U+0600..0605, U+0660..0669,
          U+066B..066C, U+06DD, U+08E2,
          U+10D30..10D39 ...)
\p{Bidi_Class: B}   \p{Bidi_Class=Paragraph_Separator} (7)
\p{Bidi_Class: BN}   \p{Bidi_Class=Boundary_Neutral} (4016)
\p{Bidi_Class: Boundary_Neutral} (Short: \p{Bc=BN}) (4016: [^\t\n
          \cK\f\r\x1c-\x7e\x85\xa0-\xac\xae-\xff],
          U+180E, U+200B..200D, U+2060..2065,
          U+206A..206F, U+FDD0..FDEF ...)
\p{Bidi_Class: Common_Separator} (Short: \p{Bc=CS}) (15: [,.\/:
          \xa0], U+060C, U+202F, U+2044, U+FE50,
          U+FE52 ...)
\p{Bidi_Class: CS}   \p{Bidi_Class=Common_Separator} (15)
\p{Bidi_Class: EN}   \p{Bidi_Class=European_Number} (168)
\p{Bidi_Class: ES}   \p{Bidi_Class=European_Separator} (12)
\p{Bidi_Class: ET}   \p{Bidi_Class=European_Terminator} (92)
\p{Bidi_Class: European_Number} (Short: \p{Bc=EN}) (168: [0-9\xb2-
          \xb3\xb9], U+06F0..06F9, U+2070,
          U+2074..2079, U+2080..2089, U+2488..249B
          ...)
\p{Bidi_Class: European_Separator} (Short: \p{Bc=ES}) (12: [+\-],
          U+207A..207B, U+208A..208B, U+2212,
          U+FB29, U+FE62..FE63 ...)
\p{Bidi_Class: European_Terminator} (Short: \p{Bc=ET}) (92: [#\$
          \%\xa2-\xa5\xb0-\xb1], U+058F,
          U+0609..060A, U+066A, U+09F2..09F3,
          U+09FB ...)
\p{Bidi_Class: First_Strong_Isolate} (Short: \p{Bc=FSI}) (1:
          U+2068)
\p{Bidi_Class: FSI}   \p{Bidi_Class=First_Strong_Isolate} (1)
\p{Bidi_Class: L}   \p{Bidi_Class=Left_To_Right} (1_096_473
          plus all above-Unicode code points)
\p{Bidi_Class: Left_To_Right} (Short: \p{Bc=L}) (1_096_473 plus
          all above-Unicode code points: [A-Za-z
          \xaa\xb5\xba\xc0-\xd6\xd8-\xf6\xf8-
          \xff], U+0100..02B8, U+02BB..02C1,

U+02D0..02D1, U+02E0..02E4, U+02EE ...)
\p{Bidi_Class: Left_To_Right_Embedding} (Short: \p{Bc=LRE}) (1:
          U+202A)
\p{Bidi_Class: Left_To_Right_Isolate} (Short: \p{Bc=LRI}) (1:
          U+2066)
\p{Bidi_Class: Left_To_Right_Override} (Short: \p{Bc=LRO}) (1:
          U+202D)
\p{Bidi_Class: LRE}   \p{Bidi_Class=Left_To_Right_Embedding} (1)
\p{Bidi_Class: LRI}   \p{Bidi_Class=Left_To_Right_Isolate} (1)
\p{Bidi_Class: LRO}   \p{Bidi_Class=Left_To_Right_Override} (1)
\p{Bidi_Class: Nonspacing_Mark} (Short: \p{Bc=NSM}) (1847:
          U+0300..036F, U+0483..0489,
          U+0591..05BD, U+05BF, U+05C1..05C2,
          U+05C4..05C5 ...)
\p{Bidi_Class: NSM}   \p{Bidi_Class=Nonspacing_Mark} (1847)
\p{Bidi_Class: ON}   \p{Bidi_Class=Other_Neutral} (5931)
\p{Bidi_Class: Other_Neutral} (Short: \p{Bc=ON}) (5931: [!\"&\'
          \(\)*;<=>?\@\[\\\]\^_'\{\|\}~\xa1\xa6-
          \xa9\xab-\xac\xae-\xaf\xb4\xb6-\xb8\xbb-
          \xbf\xd7\xf7], U+02B9..02BA,
          U+02C2..02CF, U+02D2..02DF,
          U+02E5..02ED, U+02EF..02FF ...)
\p{Bidi_Class: Paragraph_Separator} (Short: \p{Bc=B}) (7: [\n\r
          \x1c-\x1e\x85], U+2029)
\p{Bidi_Class: PDF}   \p{Bidi_Class=Pop_Directional_Format} (1)
\p{Bidi_Class: PDI}   \p{Bidi_Class=Pop_Directional_Isolate} (1)
\p{Bidi_Class: Pop_Directional_Format} (Short: \p{Bc=PDF}) (1:
          U+202C)
\p{Bidi_Class: Pop_Directional_Isolate} (Short: \p{Bc=PDI}) (1:
          U+2069)
\p{Bidi_Class: R}   \p{Bidi_Class=Right_To_Left} (3763)
\p{Bidi_Class: Right_To_Left} (Short: \p{Bc=R}) (3763: U+0590,
          U+05BE, U+05C0, U+05C3, U+05C6,
          U+05C8..05FF ...)
\p{Bidi_Class: Right_To_Left_Embedding} (Short: \p{Bc=RLE}) (1:
          U+202B)
\p{Bidi_Class: Right_To_Left_Isolate} (Short: \p{Bc=RLI}) (1:
          U+2067)
\p{Bidi_Class: Right_To_Left_Override} (Short: \p{Bc=RLO}) (1:
          U+202E)

\p{Bidi_Class: RLE}   \p{Bidi_Class=Right_To_Left_Embedding} (1)
\p{Bidi_Class: RLI}   \p{Bidi_Class=Right_To_Left_Isolate} (1)
\p{Bidi_Class: RLO}   \p{Bidi_Class=Right_To_Left_Override} (1)
\p{Bidi_Class: S}   \p{Bidi_Class=Segment_Separator} (3)
\p{Bidi_Class: Segment_Separator} (Short: \p{Bc=S}) (3: [\t\cK
                \x1f])
\p{Bidi_Class: White_Space} (Short: \p{Bc=WS}) (17: [\f\x20],
                U+1680, U+2000..200A, U+2028, U+205F,
                U+3000)
\p{Bidi_Class: WS}   \p{Bidi_Class=White_Space} (17)
\p{Bidi_Control}   \p{Bidi_Control=Y} (Short: \p{BidiC}) (12)
\p{Bidi_Control: N*}   (Short: \p{BidiC=N}, \P{BidiC}) (1_114_100
                plus all above-Unicode code points:
                U+0000..061B, U+061D..200D,
                U+2010..2029, U+202F..2065,
                U+206A..infinity)
\p{Bidi_Control: Y*}   (Short: \p{BidiC=Y}, \p{BidiC}) (12:
                U+061C, U+200E..200F, U+202A..202E,
                U+2066..2069)
\p{Bidi_M}   \p{Bidi_Mirrored} (= \p{Bidi_Mirrored=Y})
                (545)
\p{Bidi_M: *}   \p{Bidi_Mirrored: *}
\p{Bidi_Mirrored}   \p{Bidi_Mirrored=Y} (Short: \p{BidiM})
                (545)
\p{Bidi_Mirrored: N*}   (Short: \p{BidiM=N}, \P{BidiM}) (1_113_567
                plus all above-Unicode code points:
                [\x00-\x20!\"#\$\%&\'*+,\-.\/0-9:;=?\@A-
                Z\\\^_'a-z\|~\x7f-\xaa\xac-\xba\xbc-
                \xff], U+0100..0F39, U+0F3E..169A,
                U+169D..2038, U+203B..2044, U+2047..207C
                ...)
\p{Bidi_Mirrored: Y*}   (Short: \p{BidiM=Y}, \p{BidiM}) (545:
                [\(\)<>\[\]\{\}\xab\xbb], U+0F3A..0F3D,
                U+169B..169C, U+2039..203A,
                U+2045..2046, U+207D..207E ...)
\p{Bidi_Paired_Bracket_Type: C} \p{Bidi_Paired_Bracket_Type=Close}
                (60)
\p{Bidi_Paired_Bracket_Type: Close} (Short: \p{Bpt=C}) (60: [\)\]
                \}], U+0F3B, U+0F3D, U+169C, U+2046,
                U+207E ...)

\p{Bidi_Paired_Bracket_Type: N} \p{Bidi_Paired_Bracket_Type=None}
                (1_113_992 plus all above-Unicode code
                points)
\p{Bidi_Paired_Bracket_Type: None} (Short: \p{Bpt=N}) (1_113_992
                plus all above-Unicode code points:
                [\x00-\x20!\"#\$\%&\'*+,\-.\/0-9:;<=>?
                \@A-Z\\\^_'a-z\|~\x7f-\xff],
                U+0100..0F39, U+0F3E..169A,
                U+169D..2044, U+2047..207C, U+207F..208C
                ...)
\p{Bidi_Paired_Bracket_Type: O} \p{Bidi_Paired_Bracket_Type=Open}
                (60)
\p{Bidi_Paired_Bracket_Type: Open} (Short: \p{Bpt=O}) (60:
                [\(\[\{], U+0F3A, U+0F3C, U+169B,
                U+2045, U+207D ...)
\p{Blank}   \p{XPosixBlank} (18)
\p{Blk: *}   \p{Block: *}
\p{Block: Adlam}   (NOT \p{Adlam} NOR \p{Is_Adlam}) (96:
                U+1E900..1E95F)
\p{Block: Aegean_Numbers} (64: U+10100..1013F)
\p{Block: Ahom}   (NOT \p{Ahom} NOR \p{Is_Ahom}) (64:
                U+11700..1173F)
\p{Block: Alchemical}   \p{Block=Alchemical_Symbols} (128)
\p{Block: Alchemical_Symbols} (Short: \p{Blk=Alchemical}) (128:
                U+1F700..1F77F)
\p{Block: Alphabetic_PF} \p{Block=Alphabetic_Presentation_Forms}
                (80)
\p{Block: Alphabetic_Presentation_Forms} (Short: \p{Blk=
                AlphabeticPF}) (80: U+FB00..FB4F)
\p{Block: Anatolian_Hieroglyphs} (NOT \p{Anatolian_Hieroglyphs}
                NOR \p{Is_Anatolian_Hieroglyphs}) (640:
                U+14400..1467F)
\p{Block: Ancient_Greek_Music} \p{Block=
                Ancient_Greek_Musical_Notation} (80)
\p{Block: Ancient_Greek_Musical_Notation} (Short: \p{Blk=
                AncientGreekMusic}) (80: U+1D200..1D24F)
\p{Block: Ancient_Greek_Numbers} (80: U+10140..1018F)
\p{Block: Ancient_Symbols} (64: U+10190..101CF)
\p{Block: Arabic}   (NOT \p{Arabic} NOR \p{Is_Arabic}) (256:
                U+0600..06FF)

\p{Block: Arabic_Ext_A} \p{Block=Arabic_Extended_A} (96)
\p{Block: Arabic_Extended_A} (Short: \p{Blk=ArabicExtA}) (96:
           U+08A0..08FF)
\p{Block: Arabic_Math} \p{Block=
           Arabic_Mathematical_Alphabetic_Symbols}
           (256)
\p{Block: Arabic_Mathematical_Alphabetic_Symbols} (Short: \p{Blk=
           ArabicMath}) (256: U+1EE00..1EEFF)
\p{Block: Arabic_PF_A} \p{Block=Arabic_Presentation_Forms_A} (688)
\p{Block: Arabic_PF_B} \p{Block=Arabic_Presentation_Forms_B} (144)
\p{Block: Arabic_Presentation_Forms_A} (Short: \p{Blk=ArabicPFA})
           (688: U+FB50..FDFF)
\p{Block: Arabic_Presentation_Forms_B} (Short: \p{Blk=ArabicPFB})
           (144: U+FE70..FEFF)
\p{Block: Arabic_Sup} \p{Block=Arabic_Supplement} (48)
\p{Block: Arabic_Supplement} (Short: \p{Blk=ArabicSup}) (48:
           U+0750..077F)
\p{Block: Armenian} (NOT \p{Armenian} NOR \p{Is_Armenian})
           (96: U+0530..058F)
\p{Block: Arrows} (112: U+2190..21FF)
\p{Block: ASCII} \p{Block=Basic_Latin} (128)
\p{Block: Avestan} (NOT \p{Avestan} NOR \p{Is_Avestan}) (64:
           U+10B00..10B3F)
\p{Block: Balinese} (NOT \p{Balinese} NOR \p{Is_Balinese})
           (128: U+1B00..1B7F)
\p{Block: Bamum} (NOT \p{Bamum} NOR \p{Is_Bamum}) (96:
           U+A6A0..A6FF)
\p{Block: Bamum_Sup} \p{Block=Bamum_Supplement} (576)
\p{Block: Bamum_Supplement} (Short: \p{Blk=BamumSup}) (576:
           U+16800..16A3F)
\p{Block: Basic_Latin} (Short: \p{Blk=ASCII}) (128: [\x00-\x7f])
\p{Block: Bassa_Vah} (NOT \p{Bassa_Vah} NOR \p{Is_Bassa_Vah})
           (48: U+16AD0..16AFF)
\p{Block: Batak} (NOT \p{Batak} NOR \p{Is_Batak}) (64:
           U+1BC0..1BFF)
\p{Block: Bengali} (NOT \p{Bengali} NOR \p{Is_Bengali}) (128:
           U+0980..09FF)
\p{Block: Bhaiksuki} (NOT \p{Bhaiksuki} NOR \p{Is_Bhaiksuki})
           (112: U+11C00..11C6F)
\p{Block: Block_Elements} (32: U+2580..259F)

\p{Block: Bopomofo}   (NOT \p{Bopomofo} NOR \p{Is_Bopomofo})
                (48: U+3100..312F)
\p{Block: Bopomofo_Ext} \p{Block=Bopomofo_Extended} (32)
\p{Block: Bopomofo_Extended} (Short: \p{Blk=BopomofoExt}) (32:
                U+31A0..31BF)
\p{Block: Box_Drawing} (128: U+2500..257F)
\p{Block: Brahmi}   (NOT \p{Brahmi} NOR \p{Is_Brahmi}) (128:
                U+11000..1107F)
\p{Block: Braille}   \p{Block=Braille_Patterns} (256)
\p{Block: Braille_Patterns} (Short: \p{Blk=Braille}) (256:
                U+2800..28FF)
\p{Block: Buginese}   (NOT \p{Buginese} NOR \p{Is_Buginese})
                (32: U+1A00..1A1F)
\p{Block: Buhid}   (NOT \p{Buhid} NOR \p{Is_Buhid}) (32:
                U+1740..175F)
\p{Block: Byzantine_Music} \p{Block=Byzantine_Musical_Symbols}
                (256)
\p{Block: Byzantine_Musical_Symbols} (Short: \p{Blk=
                ByzantineMusic}) (256: U+1D000..1D0FF)
\p{Block: Canadian_Syllabics} \p{Block=
                Unified_Canadian_Aboriginal_Syllabics}
                (640)
\p{Block: Carian}   (NOT \p{Carian} NOR \p{Is_Carian}) (64:
                U+102A0..102DF)
\p{Block: Caucasian_Albanian} (NOT \p{Caucasian_Albanian} NOR
                \p{Is_Caucasian_Albanian}) (64:
                U+10530..1056F)
\p{Block: Chakma}   (NOT \p{Chakma} NOR \p{Is_Chakma}) (80:
                U+11100..1114F)
\p{Block: Cham}   (NOT \p{Cham} NOR \p{Is_Cham}) (96:
                U+AA00..AA5F)
\p{Block: Cherokee}   (NOT \p{Cherokee} NOR \p{Is_Cherokee})
                (96: U+13A0..13FF)
\p{Block: Cherokee_Sup} \p{Block=Cherokee_Supplement} (80)
\p{Block: Cherokee_Supplement} (Short: \p{Blk=CherokeeSup}) (80:
                U+AB70..ABBF)
\p{Block: Chess_Symbols} (112: U+1FA00..1FA6F)
\p{Block: Chorasmian}   (NOT \p{Chorasmian} NOR \p{Is_Chorasmian})
                (48: U+10FB0..10FDF)
\p{Block: CJK}   \p{Block=CJK_Unified_Ideographs} (20_992)

\p{Block: CJK_Compat} \p{Block=CJK_Compatibility} (256)
\p{Block: CJK_Compat_Forms} \p{Block=CJK_Compatibility_Forms} (32)
\p{Block: CJK_Compat_Ideographs} \p{Block=
             CJK_Compatibility_Ideographs} (512)
\p{Block: CJK_Compat_Ideographs_Sup} \p{Block=
             CJK_Compatibility_Ideographs_Supplement}
             (544)
\p{Block: CJK_Compatibility} (Short: \p{Blk=CJKCompat}) (256:
             U+3300..33FF)
\p{Block: CJK_Compatibility_Forms} (Short: \p{Blk=CJKCompatForms})
             (32: U+FE30..FE4F)
\p{Block: CJK_Compatibility_Ideographs} (Short: \p{Blk=
             CJKCompatIdeographs}) (512: U+F900..FAFF)
\p{Block: CJK_Compatibility_Ideographs_Supplement} (Short: \p{Blk=
             CJKCompatIdeographsSup}) (544:
             U+2F800..2FA1F)
\p{Block: CJK_Ext_A}   \p{Block=
             CJK_Unified_Ideographs_Extension_A}
             (6592)
\p{Block: CJK_Ext_B}   \p{Block=
             CJK_Unified_Ideographs_Extension_B}
             (42_720)
\p{Block: CJK_Ext_C}   \p{Block=
             CJK_Unified_Ideographs_Extension_C}
             (4160)
\p{Block: CJK_Ext_D}   \p{Block=
             CJK_Unified_Ideographs_Extension_D} (224)
\p{Block: CJK_Ext_E}   \p{Block=
             CJK_Unified_Ideographs_Extension_E}
             (5776)
\p{Block: CJK_Ext_F}   \p{Block=
             CJK_Unified_Ideographs_Extension_F}
             (7488)
\p{Block: CJK_Ext_G}   \p{Block=
             CJK_Unified_Ideographs_Extension_G}
             (4944)
\p{Block: CJK_Radicals_Sup} \p{Block=CJK_Radicals_Supplement} (128)
\p{Block: CJK_Radicals_Supplement} (Short: \p{Blk=CJKRadicalsSup})
             (128: U+2E80..2EFF)
\p{Block: CJK_Strokes} (48: U+31C0..31EF)

\p{Block: CJK_Symbols} \p{Block=CJK_Symbols_And_Punctuation} (64)
\p{Block: CJK_Symbols_And_Punctuation} (Short: \p{Blk=CJKSymbols})
        (64: U+3000..303F)
\p{Block: CJK_Unified_Ideographs} (Short: \p{Blk=CJK}) (20_992:
        U+4E00..9FFF)
\p{Block: CJK_Unified_Ideographs_Extension_A} (Short: \p{Blk=
        CJKExtA}) (6592: U+3400..4DBF)
\p{Block: CJK_Unified_Ideographs_Extension_B} (Short: \p{Blk=
        CJKExtB}) (42_720: U+20000..2A6DF)
\p{Block: CJK_Unified_Ideographs_Extension_C} (Short: \p{Blk=
        CJKExtC}) (4160: U+2A700..2B73F)
\p{Block: CJK_Unified_Ideographs_Extension_D} (Short: \p{Blk=
        CJKExtD}) (224: U+2B740..2B81F)
\p{Block: CJK_Unified_Ideographs_Extension_E} (Short: \p{Blk=
        CJKExtE}) (5776: U+2B820..2CEAF)
\p{Block: CJK_Unified_Ideographs_Extension_F} (Short: \p{Blk=
        CJKExtF}) (7488: U+2CEB0..2EBEF)
\p{Block: CJK_Unified_Ideographs_Extension_G} (Short: \p{Blk=
        CJKExtG}) (4944: U+30000..3134F)
\p{Block: Combining_Diacritical_Marks} (Short: \p{Blk=
        Diacriticals}) (112: U+0300..036F)
\p{Block: Combining_Diacritical_Marks_Extended} (Short: \p{Blk=
        DiacriticalsExt}) (80: U+1AB0..1AFF)
\p{Block: Combining_Diacritical_Marks_For_Symbols} (Short: \p{Blk=
        DiacriticalsForSymbols}) (48:
        U+20D0..20FF)
\p{Block: Combining_Diacritical_Marks_Supplement} (Short: \p{Blk=
        DiacriticalsSup}) (64: U+1DC0..1DFF)
\p{Block: Combining_Half_Marks} (Short: \p{Blk=HalfMarks}) (16:
        U+FE20..FE2F)
\p{Block: Combining_Marks_For_Symbols} \p{Block=
        Combining_Diacritical_Marks_For_Symbols}
        (48)
\p{Block: Common_Indic_Number_Forms} (Short: \p{Blk=
        IndicNumberForms}) (16: U+A830..A83F)
\p{Block: Compat_Jamo} \p{Block=Hangul_Compatibility_Jamo} (96)
\p{Block: Control_Pictures} (64: U+2400..243F)
\p{Block: Coptic}   (NOT \p{Coptic} NOR \p{Is_Coptic}) (128:
        U+2C80..2CFF)
\p{Block: Coptic_Epact_Numbers} (32: U+102E0..102FF)

\p{Block: Counting_Rod} \p{Block=Counting_Rod_Numerals} (32)
\p{Block: Counting_Rod_Numerals} (Short: \p{Blk=CountingRod}) (32:
          U+1D360..1D37F)
\p{Block: Cuneiform}   (NOT \p{Cuneiform} NOR \p{Is_Cuneiform})
          (1024: U+12000..123FF)
\p{Block: Cuneiform_Numbers} \p{Block=
          Cuneiform_Numbers_And_Punctuation} (128)
\p{Block: Cuneiform_Numbers_And_Punctuation} (Short: \p{Blk=
          CuneiformNumbers}) (128: U+12400..1247F)
\p{Block: Currency_Symbols} (48: U+20A0..20CF)
\p{Block: Cypriot_Syllabary} (64: U+10800..1083F)
\p{Block: Cyrillic}   (NOT \p{Cyrillic} NOR \p{Is_Cyrillic})
          (256: U+0400..04FF)
\p{Block: Cyrillic_Ext_A} \p{Block=Cyrillic_Extended_A} (32)
\p{Block: Cyrillic_Ext_B} \p{Block=Cyrillic_Extended_B} (96)
\p{Block: Cyrillic_Ext_C} \p{Block=Cyrillic_Extended_C} (16)
\p{Block: Cyrillic_Extended_A} (Short: \p{Blk=CyrillicExtA}) (32:
          U+2DE0..2DFF)
\p{Block: Cyrillic_Extended_B} (Short: \p{Blk=CyrillicExtB}) (96:
          U+A640..A69F)
\p{Block: Cyrillic_Extended_C} (Short: \p{Blk=CyrillicExtC}) (16:
          U+1C80..1C8F)
\p{Block: Cyrillic_Sup} \p{Block=Cyrillic_Supplement} (48)
\p{Block: Cyrillic_Supplement} (Short: \p{Blk=CyrillicSup}) (48:
          U+0500..052F)
\p{Block: Cyrillic_Supplementary} \p{Block=Cyrillic_Supplement}
          (48)
\p{Block: Deseret}   (80: U+10400..1044F)
\p{Block: Devanagari}   (NOT \p{Devanagari} NOR \p{Is_Devanagari})
          (128: U+0900..097F)
\p{Block: Devanagari_Ext} \p{Block=Devanagari_Extended} (32)
\p{Block: Devanagari_Extended} (Short: \p{Blk=DevanagariExt}) (32:
          U+A8E0..A8FF)
\p{Block: Diacriticals} \p{Block=Combining_Diacritical_Marks} (112)
\p{Block: Diacriticals_Ext} \p{Block=
          Combining_Diacritical_Marks_Extended}
          (80)
\p{Block: Diacriticals_For_Symbols} \p{Block=
          Combining_Diacritical_Marks_For_Symbols}
          (48)

\p{Block: Diacriticals_Sup} \p{Block=
             Combining_Diacritical_Marks_Supplement}
             (64)
\p{Block: Dingbats}   (192: U+2700..27BF)
\p{Block: Dives_Akuru} (NOT \p{Dives_Akuru} NOR
             \p{Is_Dives_Akuru}) (96: U+11900..1195F)
\p{Block: Dogra}   (NOT \p{Dogra} NOR \p{Is_Dogra}) (80:
             U+11800..1184F)
\p{Block: Domino}   \p{Block=Domino_Tiles} (112)
\p{Block: Domino_Tiles} (Short: \p{Blk=Domino}) (112:
             U+1F030..1F09F)
\p{Block: Duployan}   (NOT \p{Duployan} NOR \p{Is_Duployan})
             (160: U+1BC00..1BC9F)
\p{Block: Early_Dynastic_Cuneiform} (208: U+12480..1254F)
\p{Block: Egyptian_Hieroglyph_Format_Controls} (16: U+13430..1343F)
\p{Block: Egyptian_Hieroglyphs} (NOT \p{Egyptian_Hieroglyphs} NOR
             \p{Is_Egyptian_Hieroglyphs}) (1072:
             U+13000..1342F)
\p{Block: Elbasan}   (NOT \p{Elbasan} NOR \p{Is_Elbasan}) (48:
             U+10500..1052F)
\p{Block: Elymaic}   (NOT \p{Elymaic} NOR \p{Is_Elymaic}) (32:
             U+10FE0..10FFF)
\p{Block: Emoticons}   (80: U+1F600..1F64F)
\p{Block: Enclosed_Alphanum} \p{Block=Enclosed_Alphanumerics} (160)
\p{Block: Enclosed_Alphanum_Sup} \p{Block=
             Enclosed_Alphanumeric_Supplement} (256)
\p{Block: Enclosed_Alphanumeric_Supplement} (Short: \p{Blk=
             EnclosedAlphanumSup}) (256:
             U+1F100..1F1FF)
\p{Block: Enclosed_Alphanumerics} (Short: \p{Blk=
             EnclosedAlphanum}) (160: U+2460..24FF)
\p{Block: Enclosed_CJK} \p{Block=Enclosed_CJK_Letters_And_Months}
             (256)
\p{Block: Enclosed_CJK_Letters_And_Months} (Short: \p{Blk=
             EnclosedCJK}) (256: U+3200..32FF)
\p{Block: Enclosed_Ideographic_Sup} \p{Block=
             Enclosed_Ideographic_Supplement} (256)
\p{Block: Enclosed_Ideographic_Supplement} (Short: \p{Blk=
             EnclosedIdeographicSup}) (256:
             U+1F200..1F2FF)

\p{Block: Ethiopic}   (NOT \p{Ethiopic} NOR \p{Is_Ethiopic})
                 (384: U+1200..137F)
\p{Block: Ethiopic_Ext} \p{Block=Ethiopic_Extended} (96)
\p{Block: Ethiopic_Ext_A} \p{Block=Ethiopic_Extended_A} (48)
\p{Block: Ethiopic_Extended} (Short: \p{Blk=EthiopicExt}) (96:
                 U+2D80..2DDF)
\p{Block: Ethiopic_Extended_A} (Short: \p{Blk=EthiopicExtA}) (48:
                 U+AB00..AB2F)
\p{Block: Ethiopic_Sup} \p{Block=Ethiopic_Supplement} (32)
\p{Block: Ethiopic_Supplement} (Short: \p{Blk=EthiopicSup}) (32:
                 U+1380..139F)
\p{Block: General_Punctuation} (Short: \p{Blk=Punctuation}; NOT
                 \p{Punct} NOR \p{Is_Punctuation}) (112:
                 U+2000..206F)
\p{Block: Geometric_Shapes} (96: U+25A0..25FF)
\p{Block: Geometric_Shapes_Ext} \p{Block=
                 Geometric_Shapes_Extended} (128)
\p{Block: Geometric_Shapes_Extended} (Short: \p{Blk=
                 GeometricShapesExt}) (128:
                 U+1F780..1F7FF)
\p{Block: Georgian}   (NOT \p{Georgian} NOR \p{Is_Georgian})
                 (96: U+10A0..10FF)
\p{Block: Georgian_Ext} \p{Block=Georgian_Extended} (48)
\p{Block: Georgian_Extended} (Short: \p{Blk=GeorgianExt}) (48:
                 U+1C90..1CBF)
\p{Block: Georgian_Sup} \p{Block=Georgian_Supplement} (48)
\p{Block: Georgian_Supplement} (Short: \p{Blk=GeorgianSup}) (48:
                 U+2D00..2D2F)
\p{Block: Glagolitic}   (NOT \p{Glagolitic} NOR \p{Is_Glagolitic})
                 (96: U+2C00..2C5F)
\p{Block: Glagolitic_Sup} \p{Block=Glagolitic_Supplement} (48)
\p{Block: Glagolitic_Supplement} (Short: \p{Blk=GlagoliticSup})
                 (48: U+1E000..1E02F)
\p{Block: Gothic}   (NOT \p{Gothic} NOR \p{Is_Gothic}) (32:
                 U+10330..1034F)
\p{Block: Grantha}   (NOT \p{Grantha} NOR \p{Is_Grantha}) (128:
                 U+11300..1137F)
\p{Block: Greek}   \p{Block=Greek_And_Coptic} (NOT \p{Greek}
                 NOR \p{Is_Greek}) (144)
\p{Block: Greek_And_Coptic} (Short: \p{Blk=Greek}; NOT \p{Greek}

NOR \p{Is_Greek}) (144: U+0370..03FF)

\p{Block: Greek_Ext}   \p{Block=Greek_Extended} (256)

\p{Block: Greek_Extended} (Short: \p{Blk=GreekExt}) (256:
          U+1F00..1FFF)

\p{Block: Gujarati}   (NOT \p{Gujarati} NOR \p{Is_Gujarati})
          (128: U+0A80..0AFF)

\p{Block: Gunjala_Gondi} (NOT \p{Gunjala_Gondi} NOR
          \p{Is_Gunjala_Gondi}) (80:
          U+11D60..11DAF)

\p{Block: Gurmukhi}   (NOT \p{Gurmukhi} NOR \p{Is_Gurmukhi})
          (128: U+0A00..0A7F)

\p{Block: Half_And_Full_Forms} \p{Block=
          Halfwidth_And_Fullwidth_Forms} (240)

\p{Block: Half_Marks}   \p{Block=Combining_Half_Marks} (16)

\p{Block: Halfwidth_And_Fullwidth_Forms} (Short: \p{Blk=
          HalfAndFullForms}) (240: U+FF00..FFEF)

\p{Block: Hangul}   \p{Block=Hangul_Syllables} (NOT \p{Hangul}
          NOR \p{Is_Hangul}) (11_184)

\p{Block: Hangul_Compatibility_Jamo} (Short: \p{Blk=CompatJamo})
          (96: U+3130..318F)

\p{Block: Hangul_Jamo} (Short: \p{Blk=Jamo}) (256: U+1100..11FF)

\p{Block: Hangul_Jamo_Extended_A} (Short: \p{Blk=JamoExtA}) (32:
          U+A960..A97F)

\p{Block: Hangul_Jamo_Extended_B} (Short: \p{Blk=JamoExtB}) (80:
          U+D7B0..D7FF)

\p{Block: Hangul_Syllables} (Short: \p{Blk=Hangul}; NOT \p{Hangul}
          NOR \p{Is_Hangul}) (11_184: U+AC00..D7AF)

\p{Block: Hanifi_Rohingya} (NOT \p{Hanifi_Rohingya} NOR
          \p{Is_Hanifi_Rohingya}) (64:
          U+10D00..10D3F)

\p{Block: Hanunoo}   (NOT \p{Hanunoo} NOR \p{Is_Hanunoo}) (32:
          U+1720..173F)

\p{Block: Hatran}   (NOT \p{Hatran} NOR \p{Is_Hatran}) (32:
          U+108E0..108FF)

\p{Block: Hebrew}   (NOT \p{Hebrew} NOR \p{Is_Hebrew}) (112:
          U+0590..05FF)

\p{Block: High_Private_Use_Surrogates} (Short: \p{Blk=
          HighPUSurrogates}) (128: U+DB80..DBFF)

\p{Block: High_PU_Surrogates} \p{Block=
          High_Private_Use_Surrogates} (128)

\p{Block: High_Surrogates} (896: U+D800..DB7F)
\p{Block: Hiragana}   (NOT \p{Hiragana} NOR \p{Is_Hiragana})
               (96: U+3040..309F)
\p{Block: IDC}   \p{Block=
               Ideographic_Description_Characters} (NOT
               \p{ID_Continue} NOR \p{Is_IDC}) (16)
\p{Block: Ideographic_Description_Characters} (Short: \p{Blk=IDC};
               NOT \p{ID_Continue} NOR \p{Is_IDC}) (16:
               U+2FF0..2FFF)
\p{Block: Ideographic_Symbols} \p{Block=
               Ideographic_Symbols_And_Punctuation} (32)
\p{Block: Ideographic_Symbols_And_Punctuation} (Short: \p{Blk=
               IdeographicSymbols}) (32: U+16FE0..16FFF)
\p{Block: Imperial_Aramaic} (NOT \p{Imperial_Aramaic} NOR
               \p{Is_Imperial_Aramaic}) (32:
               U+10840..1085F)
\p{Block: Indic_Number_Forms} \p{Block=Common_Indic_Number_Forms}
               (16)
\p{Block: Indic_Siyaq_Numbers} (80: U+1EC70..1ECBF)
\p{Block: Inscriptional_Pahlavi} (NOT \p{Inscriptional_Pahlavi}
               NOR \p{Is_Inscriptional_Pahlavi}) (32:
               U+10B60..10B7F)
\p{Block: Inscriptional_Parthian} (NOT \p{Inscriptional_Parthian}
               NOR \p{Is_Inscriptional_Parthian}) (32:
               U+10B40..10B5F)
\p{Block: IPA_Ext}   \p{Block=IPA_Extensions} (96)
\p{Block: IPA_Extensions} (Short: \p{Blk=IPAExt}) (96:
               U+0250..02AF)
\p{Block: Jamo}   \p{Block=Hangul_Jamo} (256)
\p{Block: Jamo_Ext_A}   \p{Block=Hangul_Jamo_Extended_A} (32)
\p{Block: Jamo_Ext_B}   \p{Block=Hangul_Jamo_Extended_B} (80)
\p{Block: Javanese}   (NOT \p{Javanese} NOR \p{Is_Javanese})
               (96: U+A980..A9DF)
\p{Block: Kaithi}   (NOT \p{Kaithi} NOR \p{Is_Kaithi}) (80:
               U+11080..110CF)
\p{Block: Kana_Ext_A}   \p{Block=Kana_Extended_A} (48)
\p{Block: Kana_Extended_A} (Short: \p{Blk=KanaExtA}) (48:
               U+1B100..1B12F)
\p{Block: Kana_Sup}   \p{Block=Kana_Supplement} (256)
\p{Block: Kana_Supplement} (Short: \p{Blk=KanaSup}) (256:

                    U+1B000..1B0FF)
    \p{Block: Kanbun}   (16: U+3190..319F)
    \p{Block: Kangxi}   \p{Block=Kangxi_Radicals} (224)
    \p{Block: Kangxi_Radicals} (Short: \p{Blk=Kangxi}) (224:
                    U+2F00..2FDF)
    \p{Block: Kannada}   (NOT \p{Kannada} NOR \p{Is_Kannada}) (128:
                    U+0C80..0CFF)
    \p{Block: Katakana}   (NOT \p{Katakana} NOR \p{Is_Katakana})
                    (96: U+30A0..30FF)
    \p{Block: Katakana_Ext} \p{Block=Katakana_Phonetic_Extensions} (16)
    \p{Block: Katakana_Phonetic_Extensions} (Short: \p{Blk=
                    KatakanaExt}) (16: U+31F0..31FF)
    \p{Block: Kayah_Li}   (48: U+A900..A92F)
    \p{Block: Kharoshthi}   (NOT \p{Kharoshthi} NOR \p{Is_Kharoshthi})
                    (96: U+10A00..10A5F)
    \p{Block: Khitan_Small_Script} (NOT \p{Khitan_Small_Script} NOR
                    \p{Is_Khitan_Small_Script}) (512:
                    U+18B00..18CFF)
    \p{Block: Khmer}   (NOT \p{Khmer} NOR \p{Is_Khmer}) (128:
                    U+1780..17FF)
    \p{Block: Khmer_Symbols} (32: U+19E0..19FF)
    \p{Block: Khojki}   (NOT \p{Khojki} NOR \p{Is_Khojki}) (80:
                    U+11200..1124F)
    \p{Block: Khudawadi}   (NOT \p{Khudawadi} NOR \p{Is_Khudawadi})
                    (80: U+112B0..112FF)
    \p{Block: Lao}   (NOT \p{Lao} NOR \p{Is_Lao}) (128:
                    U+0E80..0EFF)
    \p{Block: Latin_1}   \p{Block=Latin_1_Supplement} (128)
    \p{Block: Latin_1_Sup} \p{Block=Latin_1_Supplement} (128)
    \p{Block: Latin_1_Supplement} (Short: \p{Blk=Latin1}) (128: [\x80-
                    \xff])
    \p{Block: Latin_Ext_A} \p{Block=Latin_Extended_A} (128)
    \p{Block: Latin_Ext_Additional} \p{Block=
                    Latin_Extended_Additional} (256)
    \p{Block: Latin_Ext_B} \p{Block=Latin_Extended_B} (208)
    \p{Block: Latin_Ext_C} \p{Block=Latin_Extended_C} (32)
    \p{Block: Latin_Ext_D} \p{Block=Latin_Extended_D} (224)
    \p{Block: Latin_Ext_E} \p{Block=Latin_Extended_E} (64)
    \p{Block: Latin_Extended_A} (Short: \p{Blk=LatinExtA}) (128:
                    U+0100..017F)

\p{Block: Latin_Extended_Additional} (Short: \p{Blk=
             LatinExtAdditional}) (256: U+1E00..1EFF)
\p{Block: Latin_Extended_B} (Short: \p{Blk=LatinExtB}) (208:
             U+0180..024F)
\p{Block: Latin_Extended_C} (Short: \p{Blk=LatinExtC}) (32:
             U+2C60..2C7F)
\p{Block: Latin_Extended_D} (Short: \p{Blk=LatinExtD}) (224:
             U+A720..A7FF)
\p{Block: Latin_Extended_E} (Short: \p{Blk=LatinExtE}) (64:
             U+AB30..AB6F)
\p{Block: Lepcha}   (NOT \p{Lepcha} NOR \p{Is_Lepcha}) (80:
             U+1C00..1C4F)
\p{Block: Letterlike_Symbols} (80: U+2100..214F)
\p{Block: Limbu}   (NOT \p{Limbu} NOR \p{Is_Limbu}) (80:
             U+1900..194F)
\p{Block: Linear_A}   (NOT \p{Linear_A} NOR \p{Is_Linear_A})
             (384: U+10600..1077F)
\p{Block: Linear_B_Ideograms} (128: U+10080..100FF)
\p{Block: Linear_B_Syllabary} (128: U+10000..1007F)
\p{Block: Lisu}   (NOT \p{Lisu} NOR \p{Is_Lisu}) (48:
             U+A4D0..A4FF)
\p{Block: Lisu_Sup}   \p{Block=Lisu_Supplement} (16)
\p{Block: Lisu_Supplement} (Short: \p{Blk=LisuSup}) (16:
             U+11FB0..11FBF)
\p{Block: Low_Surrogates} (1024: U+DC00..DFFF)
\p{Block: Lycian}   (NOT \p{Lycian} NOR \p{Is_Lycian}) (32:
             U+10280..1029F)
\p{Block: Lydian}   (NOT \p{Lydian} NOR \p{Is_Lydian}) (32:
             U+10920..1093F)
\p{Block: Mahajani}   (NOT \p{Mahajani} NOR \p{Is_Mahajani})
             (48: U+11150..1117F)
\p{Block: Mahjong}   \p{Block=Mahjong_Tiles} (48)
\p{Block: Mahjong_Tiles} (Short: \p{Blk=Mahjong}) (48:
             U+1F000..1F02F)
\p{Block: Makasar}   (NOT \p{Makasar} NOR \p{Is_Makasar}) (32:
             U+11EE0..11EFF)
\p{Block: Malayalam}   (NOT \p{Malayalam} NOR \p{Is_Malayalam})
             (128: U+0D00..0D7F)
\p{Block: Mandaic}   (NOT \p{Mandaic} NOR \p{Is_Mandaic}) (32:
             U+0840..085F)

\p{Block: Manichaean} (NOT \p{Manichaean} NOR \p{Is_Manichaean})
          (64: U+10AC0..10AFF)
\p{Block: Marchen}   (NOT \p{Marchen} NOR \p{Is_Marchen}) (80:
          U+11C70..11CBF)
\p{Block: Masaram_Gondi} (NOT \p{Masaram_Gondi} NOR
          \p{Is_Masaram_Gondi}) (96:
          U+11D00..11D5F)
\p{Block: Math_Alphanum} \p{Block=
          Mathematical_Alphanumeric_Symbols} (1024)
\p{Block: Math_Operators} \p{Block=Mathematical_Operators} (256)
\p{Block: Mathematical_Alphanumeric_Symbols} (Short: \p{Blk=
          MathAlphanum}) (1024: U+1D400..1D7FF)
\p{Block: Mathematical_Operators} (Short: \p{Blk=MathOperators})
          (256: U+2200..22FF)
\p{Block: Mayan_Numerals} (32: U+1D2E0..1D2FF)
\p{Block: Medefaidrin} (NOT \p{Medefaidrin} NOR
          \p{Is_Medefaidrin}) (96: U+16E40..16E9F)
\p{Block: Meetei_Mayek} (NOT \p{Meetei_Mayek} NOR
          \p{Is_Meetei_Mayek}) (64: U+ABC0..ABFF)
\p{Block: Meetei_Mayek_Ext} \p{Block=Meetei_Mayek_Extensions} (32)
\p{Block: Meetei_Mayek_Extensions} (Short: \p{Blk=MeeteiMayekExt})
          (32: U+AAE0..AAFF)
\p{Block: Mende_Kikakui} (NOT \p{Mende_Kikakui} NOR
          \p{Is_Mende_Kikakui}) (224:
          U+1E800..1E8DF)
\p{Block: Meroitic_Cursive} (NOT \p{Meroitic_Cursive} NOR
          \p{Is_Meroitic_Cursive}) (96:
          U+109A0..109FF)
\p{Block: Meroitic_Hieroglyphs} (32: U+10980..1099F)
\p{Block: Miao}   (NOT \p{Miao} NOR \p{Is_Miao}) (160:
          U+16F00..16F9F)
\p{Block: Misc_Arrows} \p{Block=Miscellaneous_Symbols_And_Arrows}
          (256)
\p{Block: Misc_Math_Symbols_A} \p{Block=
          Miscellaneous_Mathematical_Symbols_A}
          (48)
\p{Block: Misc_Math_Symbols_B} \p{Block=
          Miscellaneous_Mathematical_Symbols_B}
          (128)
\p{Block: Misc_Pictographs} \p{Block=

Miscellaneous_Symbols_And_Pictographs}
(768)
\p{Block: Misc_Symbols} \p{Block=Miscellaneous_Symbols} (256)
\p{Block: Misc_Technical} \p{Block=Miscellaneous_Technical} (256)
\p{Block: Miscellaneous_Mathematical_Symbols_A} (Short: \p{Blk=
MiscMathSymbolsA}) (48: U+27C0..27EF)
\p{Block: Miscellaneous_Mathematical_Symbols_B} (Short: \p{Blk=
MiscMathSymbolsB}) (128: U+2980..29FF)
\p{Block: Miscellaneous_Symbols} (Short: \p{Blk=MiscSymbols})
(256: U+2600..26FF)
\p{Block: Miscellaneous_Symbols_And_Arrows} (Short: \p{Blk=
MiscArrows}) (256: U+2B00..2BFF)
\p{Block: Miscellaneous_Symbols_And_Pictographs} (Short: \p{Blk=
MiscPictographs}) (768: U+1F300..1F5FF)
\p{Block: Miscellaneous_Technical} (Short: \p{Blk=MiscTechnical})
(256: U+2300..23FF)
\p{Block: Modi}   (NOT \p{Modi} NOR \p{Is_Modi}) (96:
U+11600..1165F)
\p{Block: Modifier_Letters} \p{Block=Spacing_Modifier_Letters} (80)
\p{Block: Modifier_Tone_Letters} (32: U+A700..A71F)
\p{Block: Mongolian}   (NOT \p{Mongolian} NOR \p{Is_Mongolian})
(176: U+1800..18AF)
\p{Block: Mongolian_Sup} \p{Block=Mongolian_Supplement} (32)
\p{Block: Mongolian_Supplement} (Short: \p{Blk=MongolianSup}) (32:
U+11660..1167F)
\p{Block: Mro}   (NOT \p{Mro} NOR \p{Is_Mro}) (48:
U+16A40..16A6F)
\p{Block: Multani}   (NOT \p{Multani} NOR \p{Is_Multani}) (48:
U+11280..112AF)
\p{Block: Music}   \p{Block=Musical_Symbols} (256)
\p{Block: Musical_Symbols} (Short: \p{Blk=Music}) (256:
U+1D100..1D1FF)
\p{Block: Myanmar}   (NOT \p{Myanmar} NOR \p{Is_Myanmar}) (160:
U+1000..109F)
\p{Block: Myanmar_Ext_A} \p{Block=Myanmar_Extended_A} (32)
\p{Block: Myanmar_Ext_B} \p{Block=Myanmar_Extended_B} (32)
\p{Block: Myanmar_Extended_A} (Short: \p{Blk=MyanmarExtA}) (32:
U+AA60..AA7F)
\p{Block: Myanmar_Extended_B} (Short: \p{Blk=MyanmarExtB}) (32:
U+A9E0..A9FF)

\p{Block: Nabataean}   (NOT \p{Nabataean} NOR \p{Is_Nabataean})
                (48: U+10880..108AF)
\p{Block: Nandinagari} (NOT \p{Nandinagari} NOR
                \p{Is_Nandinagari}) (96: U+119A0..119FF)
\p{Block: NB}   \p{Block=No_Block} (826_640 plus all
                above-Unicode code points)
\p{Block: New_Tai_Lue} (NOT \p{New_Tai_Lue} NOR
                \p{Is_New_Tai_Lue}) (96: U+1980..19DF)
\p{Block: Newa}   (NOT \p{Newa} NOR \p{Is_Newa}) (128:
                U+11400..1147F)
\p{Block: NKo}   (NOT \p{Nko} NOR \p{Is_NKo}) (64:
                U+07C0..07FF)
\p{Block: No_Block}   (Short: \p{Blk=NB}) (826_640 plus all
                above-Unicode code points: U+0870..089F,
                U+2FE0..2FEF, U+10200..1027F,
                U+103E0..103FF, U+10570..105FF,
                U+10780..107FF ...)
\p{Block: Number_Forms} (64: U+2150..218F)
\p{Block: Nushu}   (NOT \p{Nushu} NOR \p{Is_Nushu}) (400:
                U+1B170..1B2FF)
\p{Block: Nyiakeng_Puachue_Hmong} (NOT \p{Nyiakeng_Puachue_Hmong}
                NOR \p{Is_Nyiakeng_Puachue_Hmong}) (80:
                U+1E100..1E14F)
\p{Block: OCR}   \p{Block=Optical_Character_Recognition}
                (32)
\p{Block: Ogham}   (NOT \p{Ogham} NOR \p{Is_Ogham}) (32:
                U+1680..169F)
\p{Block: Ol_Chiki}   (48: U+1C50..1C7F)
\p{Block: Old_Hungarian} (NOT \p{Old_Hungarian} NOR
                \p{Is_Old_Hungarian}) (128:
                U+10C80..10CFF)
\p{Block: Old_Italic}   (NOT \p{Old_Italic} NOR \p{Is_Old_Italic})
                (48: U+10300..1032F)
\p{Block: Old_North_Arabian} (32: U+10A80..10A9F)
\p{Block: Old_Permic}   (NOT \p{Old_Permic} NOR \p{Is_Old_Permic})
                (48: U+10350..1037F)
\p{Block: Old_Persian} (NOT \p{Old_Persian} NOR
                \p{Is_Old_Persian}) (64: U+103A0..103DF)
\p{Block: Old_Sogdian} (NOT \p{Old_Sogdian} NOR
                \p{Is_Old_Sogdian}) (48: U+10F00..10F2F)

\p{Block: Old_South_Arabian} (32: U+10A60..10A7F)
\p{Block: Old_Turkic}   (NOT \p{Old_Turkic} NOR \p{Is_Old_Turkic})
            (80: U+10C00..10C4F)
\p{Block: Optical_Character_Recognition} (Short: \p{Blk=OCR}) (32:
            U+2440..245F)
\p{Block: Oriya}   (NOT \p{Oriya} NOR \p{Is_Oriya}) (128:
            U+0B00..0B7F)
\p{Block: Ornamental_Dingbats} (48: U+1F650..1F67F)
\p{Block: Osage}   (NOT \p{Osage} NOR \p{Is_Osage}) (80:
            U+104B0..104FF)
\p{Block: Osmanya}   (NOT \p{Osmanya} NOR \p{Is_Osmanya}) (48:
            U+10480..104AF)
\p{Block: Ottoman_Siyaq_Numbers} (80: U+1ED00..1ED4F)
\p{Block: Pahawh_Hmong} (NOT \p{Pahawh_Hmong} NOR
            \p{Is_Pahawh_Hmong}) (144:
            U+16B00..16B8F)
\p{Block: Palmyrene}   (32: U+10860..1087F)
\p{Block: Pau_Cin_Hau} (NOT \p{Pau_Cin_Hau} NOR
            \p{Is_Pau_Cin_Hau}) (64: U+11AC0..11AFF)
\p{Block: Phags_Pa}   (NOT \p{Phags_Pa} NOR \p{Is_Phags_Pa})
            (64: U+A840..A87F)
\p{Block: Phaistos}   \p{Block=Phaistos_Disc} (48)
\p{Block: Phaistos_Disc} (Short: \p{Blk=Phaistos}) (48:
            U+101D0..101FF)
\p{Block: Phoenician}   (NOT \p{Phoenician} NOR \p{Is_Phoenician})
            (32: U+10900..1091F)
\p{Block: Phonetic_Ext} \p{Block=Phonetic_Extensions} (128)
\p{Block: Phonetic_Ext_Sup} \p{Block=
            Phonetic_Extensions_Supplement} (64)
\p{Block: Phonetic_Extensions} (Short: \p{Blk=PhoneticExt}) (128:
            U+1D00..1D7F)
\p{Block: Phonetic_Extensions_Supplement} (Short: \p{Blk=
            PhoneticExtSup}) (64: U+1D80..1DBF)
\p{Block: Playing_Cards} (96: U+1F0A0..1F0FF)
\p{Block: Private_Use} \p{Block=Private_Use_Area} (NOT
            \p{Private_Use} NOR \p{Is_Private_Use})
            (6400)
\p{Block: Private_Use_Area} (Short: \p{Blk=PUA}; NOT
            \p{Private_Use} NOR \p{Is_Private_Use})
            (6400: U+E000..F8FF)

\p{Block: Psalter_Pahlavi} (NOT \p{Psalter_Pahlavi} NOR
                \p{Is_Psalter_Pahlavi}) (48:
                U+10B80..10BAF)
\p{Block: PUA}   \p{Block=Private_Use_Area} (NOT
                \p{Private_Use} NOR \p{Is_Private_Use})
                (6400)
\p{Block: Punctuation} \p{Block=General_Punctuation} (NOT
                \p{Punct} NOR \p{Is_Punctuation}) (112)
\p{Block: Rejang}   (NOT \p{Rejang} NOR \p{Is_Rejang}) (48:
                U+A930..A95F)
\p{Block: Rumi}   \p{Block=Rumi_Numeral_Symbols} (32)
\p{Block: Rumi_Numeral_Symbols} (Short: \p{Blk=Rumi}) (32:
                U+10E60..10E7F)
\p{Block: Runic}   (NOT \p{Runic} NOR \p{Is_Runic}) (96:
                U+16A0..16FF)
\p{Block: Samaritan}   (NOT \p{Samaritan} NOR \p{Is_Samaritan})
                (64: U+0800..083F)
\p{Block: Saurashtra}   (NOT \p{Saurashtra} NOR \p{Is_Saurashtra})
                (96: U+A880..A8DF)
\p{Block: Sharada}   (NOT \p{Sharada} NOR \p{Is_Sharada}) (96:
                U+11180..111DF)
\p{Block: Shavian}   (48: U+10450..1047F)
\p{Block: Shorthand_Format_Controls} (16: U+1BCA0..1BCAF)
\p{Block: Siddham}   (NOT \p{Siddham} NOR \p{Is_Siddham}) (128:
                U+11580..115FF)
\p{Block: Sinhala}   (NOT \p{Sinhala} NOR \p{Is_Sinhala}) (128:
                U+0D80..0DFF)
\p{Block: Sinhala_Archaic_Numbers} (32: U+111E0..111FF)
\p{Block: Small_Form_Variants} (Short: \p{Blk=SmallForms}) (32:
                U+FE50..FE6F)
\p{Block: Small_Forms} \p{Block=Small_Form_Variants} (32)
\p{Block: Small_Kana_Ext} \p{Block=Small_Kana_Extension} (64)
\p{Block: Small_Kana_Extension} (Short: \p{Blk=SmallKanaExt}) (64:
                U+1B130..1B16F)
\p{Block: Sogdian}   (NOT \p{Sogdian} NOR \p{Is_Sogdian}) (64:
                U+10F30..10F6F)
\p{Block: Sora_Sompeng} (NOT \p{Sora_Sompeng} NOR
                \p{Is_Sora_Sompeng}) (48: U+110D0..110FF)
\p{Block: Soyombo}   (NOT \p{Soyombo} NOR \p{Is_Soyombo}) (96:
                U+11A50..11AAF)

\p{Block: Spacing_Modifier_Letters} (Short: \p{Blk=
           ModifierLetters}) (80: U+02B0..02FF)

\p{Block: Specials}   (16: U+FFF0..FFFF)

\p{Block: Sundanese}   (NOT \p{Sundanese} NOR \p{Is_Sundanese})
           (64: U+1B80..1BBF)

\p{Block: Sundanese_Sup} \p{Block=Sundanese_Supplement} (16)

\p{Block: Sundanese_Supplement} (Short: \p{Blk=SundaneseSup}) (16:
           U+1CC0..1CCF)

\p{Block: Sup_Arrows_A} \p{Block=Supplemental_Arrows_A} (16)

\p{Block: Sup_Arrows_B} \p{Block=Supplemental_Arrows_B} (128)

\p{Block: Sup_Arrows_C} \p{Block=Supplemental_Arrows_C} (256)

\p{Block: Sup_Math_Operators} \p{Block=
           Supplemental_Mathematical_Operators}
           (256)

\p{Block: Sup_PUA_A}   \p{Block=Supplementary_Private_Use_Area_A}
           (65_536)

\p{Block: Sup_PUA_B}   \p{Block=Supplementary_Private_Use_Area_B}
           (65_536)

\p{Block: Sup_Punctuation} \p{Block=Supplemental_Punctuation} (128)

\p{Block: Sup_Symbols_And_Pictographs} \p{Block=
           Supplemental_Symbols_And_Pictographs}
           (256)

\p{Block: Super_And_Sub} \p{Block=Superscripts_And_Subscripts} (48)

\p{Block: Superscripts_And_Subscripts} (Short: \p{Blk=
           SuperAndSub}) (48: U+2070..209F)

\p{Block: Supplemental_Arrows_A} (Short: \p{Blk=SupArrowsA}) (16:
           U+27F0..27FF)

\p{Block: Supplemental_Arrows_B} (Short: \p{Blk=SupArrowsB}) (128:
           U+2900..297F)

\p{Block: Supplemental_Arrows_C} (Short: \p{Blk=SupArrowsC}) (256:
           U+1F800..1F8FF)

\p{Block: Supplemental_Mathematical_Operators} (Short: \p{Blk=
           SupMathOperators}) (256: U+2A00..2AFF)

\p{Block: Supplemental_Punctuation} (Short: \p{Blk=
           SupPunctuation}) (128: U+2E00..2E7F)

\p{Block: Supplemental_Symbols_And_Pictographs} (Short: \p{Blk=
           SupSymbolsAndPictographs}) (256:
           U+1F900..1F9FF)

\p{Block: Supplementary_Private_Use_Area_A} (Short: \p{Blk=
           SupPUAA}) (65_536: U+F0000..FFFFF)

\p{Block: Supplementary_Private_Use_Area_B} (Short: \p{Blk=
          SupPUAB}) (65_536: U+100000..10FFFF)

\p{Block: Sutton_SignWriting} (688: U+1D800..1DAAF)

\p{Block: Syloti_Nagri} (NOT \p{Syloti_Nagri} NOR
          \p{Is_Syloti_Nagri}) (48: U+A800..A82F)

\p{Block: Symbols_And_Pictographs_Ext_A} \p{Block=
          Symbols_And_Pictographs_Extended_A} (144)

\p{Block: Symbols_And_Pictographs_Extended_A} (Short: \p{Blk=
          SymbolsAndPictographsExtA}) (144:
          U+1FA70..1FAFF)

\p{Block: Symbols_For_Legacy_Computing} (256: U+1FB00..1FBFF)

\p{Block: Syriac}   (NOT \p{Syriac} NOR \p{Is_Syriac}) (80:
          U+0700..074F)

\p{Block: Syriac_Sup}   \p{Block=Syriac_Supplement} (16)

\p{Block: Syriac_Supplement} (Short: \p{Blk=SyriacSup}) (16:
          U+0860..086F)

\p{Block: Tagalog}   (NOT \p{Tagalog} NOR \p{Is_Tagalog}) (32:
          U+1700..171F)

\p{Block: Tagbanwa}   (NOT \p{Tagbanwa} NOR \p{Is_Tagbanwa})
          (32: U+1760..177F)

\p{Block: Tags}   (128: U+E0000..E007F)

\p{Block: Tai_Le}   (NOT \p{Tai_Le} NOR \p{Is_Tai_Le}) (48:
          U+1950..197F)

\p{Block: Tai_Tham}   (NOT \p{Tai_Tham} NOR \p{Is_Tai_Tham})
          (144: U+1A20..1AAF)

\p{Block: Tai_Viet}   (NOT \p{Tai_Viet} NOR \p{Is_Tai_Viet})
          (96: U+AA80..AADF)

\p{Block: Tai_Xuan_Jing} \p{Block=Tai_Xuan_Jing_Symbols} (96)

\p{Block: Tai_Xuan_Jing_Symbols} (Short: \p{Blk=TaiXuanJing}) (96:
          U+1D300..1D35F)

\p{Block: Takri}   (NOT \p{Takri} NOR \p{Is_Takri}) (80:
          U+11680..116CF)

\p{Block: Tamil}   (NOT \p{Tamil} NOR \p{Is_Tamil}) (128:
          U+0B80..0BFF)

\p{Block: Tamil_Sup}   \p{Block=Tamil_Supplement} (64)

\p{Block: Tamil_Supplement} (Short: \p{Blk=TamilSup}) (64:
          U+11FC0..11FFF)

\p{Block: Tangut}   (NOT \p{Tangut} NOR \p{Is_Tangut}) (6144:
          U+17000..187FF)

\p{Block: Tangut_Components} (768: U+18800..18AFF)

\p{Block: Tangut_Sup}   \p{Block=Tangut_Supplement} (144)
\p{Block: Tangut_Supplement} (Short: \p{Blk=TangutSup}) (144:
              U+18D00..18D8F)
\p{Block: Telugu}   (NOT \p{Telugu} NOR \p{Is_Telugu}) (128:
              U+0C00..0C7F)
\p{Block: Thaana}   (NOT \p{Thaana} NOR \p{Is_Thaana}) (64:
              U+0780..07BF)
\p{Block: Thai}   (NOT \p{Thai} NOR \p{Is_Thai}) (128:
              U+0E00..0E7F)
\p{Block: Tibetan}   (NOT \p{Tibetan} NOR \p{Is_Tibetan}) (256:
              U+0F00..0FFF)
\p{Block: Tifinagh}   (NOT \p{Tifinagh} NOR \p{Is_Tifinagh})
              (80: U+2D30..2D7F)
\p{Block: Tirhuta}   (NOT \p{Tirhuta} NOR \p{Is_Tirhuta}) (96:
              U+11480..114DF)
\p{Block: Transport_And_Map} \p{Block=Transport_And_Map_Symbols}
              (128)
\p{Block: Transport_And_Map_Symbols} (Short: \p{Blk=
              TransportAndMap}) (128: U+1F680..1F6FF)
\p{Block: UCAS}   \p{Block=
              Unified_Canadian_Aboriginal_Syllabics}
              (640)
\p{Block: UCAS_Ext}   \p{Block=
              Unified_Canadian_Aboriginal_Syllabics_-
              Extended} (80)
\p{Block: Ugaritic}   (NOT \p{Ugaritic} NOR \p{Is_Ugaritic})
              (32: U+10380..1039F)
\p{Block: Unified_Canadian_Aboriginal_Syllabics} (Short: \p{Blk=
              UCAS}) (640: U+1400..167F)
\p{Block: Unified_Canadian_Aboriginal_Syllabics_Extended} (Short:
              \p{Blk=UCASExt}) (80: U+18B0..18FF)
\p{Block: Vai}   (NOT \p{Vai} NOR \p{Is_Vai}) (320:
              U+A500..A63F)
\p{Block: Variation_Selectors} (Short: \p{Blk=VS}; NOT
              \p{Variation_Selector} NOR \p{Is_VS})
              (16: U+FE00..FE0F)
\p{Block: Variation_Selectors_Supplement} (Short: \p{Blk=VSSup})
              (240: U+E0100..E01EF)
\p{Block: Vedic_Ext}   \p{Block=Vedic_Extensions} (48)
\p{Block: Vedic_Extensions} (Short: \p{Blk=VedicExt}) (48:

```
                        U+1CD0..1CFF)
     \p{Block: Vertical_Forms} (16: U+FE10..FE1F)
     \p{Block: VS}          \p{Block=Variation_Selectors} (NOT
                       \p{Variation_Selector} NOR \p{Is_VS})
                       (16)
     \p{Block: VS_Sup}      \p{Block=Variation_Selectors_Supplement}
                       (240)
     \p{Block: Wancho}      (NOT \p{Wancho} NOR \p{Is_Wancho}) (64:
                       U+1E2C0..1E2FF)
     \p{Block: Warang_Citi}  (NOT \p{Warang_Citi} NOR
                       \p{Is_Warang_Citi}) (96: U+118A0..118FF)
     \p{Block: Yezidi}       (NOT \p{Yezidi} NOR \p{Is_Yezidi}) (64:
                       U+10E80..10EBF)
     \p{Block: Yi_Radicals}  (64: U+A490..A4CF)
     \p{Block: Yi_Syllables} (1168: U+A000..A48F)
     \p{Block: Yijing}       \p{Block=Yijing_Hexagram_Symbols} (64)
     \p{Block: Yijing_Hexagram_Symbols} (Short: \p{Blk=Yijing}) (64:
                       U+4DC0..4DFF)
     \p{Block: Zanabazar_Square} (NOT \p{Zanabazar_Square} NOR
                       \p{Is_Zanabazar_Square}) (80:
                       U+11A00..11A4F)
   X \p{Block_Elements}     \p{Block=Block_Elements} (32)
    \p{Bopo}              \p{Bopomofo} (= \p{Script_Extensions=
                       Bopomofo}) (NOT \p{Block=Bopomofo}) (117)
    \p{Bopomofo}          \p{Script_Extensions=Bopomofo} (Short:
                       \p{Bopo}; NOT \p{Block=Bopomofo}) (117)
   X \p{Bopomofo_Ext}      \p{Bopomofo_Extended} (= \p{Block=
                       Bopomofo_Extended}) (32)
   X \p{Bopomofo_Extended}  \p{Block=Bopomofo_Extended} (Short:
                       \p{InBopomofoExt}) (32)
   X \p{Box_Drawing}       \p{Block=Box_Drawing} (128)
    \p{Bpt: *}           \p{Bidi_Paired_Bracket_Type: *}
    \p{Brah}             \p{Brahmi} (= \p{Script_Extensions=
                       Brahmi}) (NOT \p{Block=Brahmi}) (109)
    \p{Brahmi}            \p{Script_Extensions=Brahmi} (Short:
                       \p{Brah}; NOT \p{Block=Brahmi}) (109)
    \p{Brai}             \p{Braille} (= \p{Script_Extensions=
                       Braille}) (256)
    \p{Braille}           \p{Script_Extensions=Braille} (Short:
                       \p{Brai}) (256)
```

X \p{Braille_Patterns}   \p{Block=Braille_Patterns} (Short:
                \p{InBraille}) (256)
 \p{Bugi}   \p{Buginese} (= \p{Script_Extensions=
                Buginese}) (NOT \p{Block=Buginese}) (31)
 \p{Buginese}   \p{Script_Extensions=Buginese} (Short:
                \p{Bugi}; NOT \p{Block=Buginese}) (31)
 \p{Buhd}   \p{Buhid} (= \p{Script_Extensions=Buhid})
                (NOT \p{Block=Buhid}) (22)
 \p{Buhid}   \p{Script_Extensions=Buhid} (Short:
                \p{Buhd}; NOT \p{Block=Buhid}) (22)
X \p{Byzantine_Music}   \p{Byzantine_Musical_Symbols} (= \p{Block=
                Byzantine_Musical_Symbols}) (256)
X \p{Byzantine_Musical_Symbols} \p{Block=Byzantine_Musical_Symbols}
                (Short: \p{InByzantineMusic}) (256)
 \p{C} \pC   \p{Other} (= \p{General_Category=Other})
                (970_414 plus all above-Unicode code
                points)
 \p{Cakm}   \p{Chakma} (= \p{Script_Extensions=
                Chakma}) (NOT \p{Block=Chakma}) (91)
 \p{Canadian_Aboriginal} \p{Script_Extensions=Canadian_Aboriginal}
                (Short: \p{Cans}) (710)
X \p{Canadian_Syllabics} \p{Unified_Canadian_Aboriginal_Syllabics}
                (= \p{Block=
                Unified_Canadian_Aboriginal_Syllabics})
                (640)
T \p{Canonical_Combining_Class: 0} \p{Canonical_Combining_Class=
                Not_Reordered} (1_113_240 plus all
                above-Unicode code points)
T \p{Canonical_Combining_Class: 1} \p{Canonical_Combining_Class=
                Overlay} (32)
T \p{Canonical_Combining_Class: 6} \p{Canonical_Combining_Class=
                Han_Reading} (2)
T \p{Canonical_Combining_Class: 7} \p{Canonical_Combining_Class=
                Nukta} (26)
T \p{Canonical_Combining_Class: 8} \p{Canonical_Combining_Class=
                Kana_Voicing} (2)
T \p{Canonical_Combining_Class: 9} \p{Canonical_Combining_Class=
                Virama} (61)
T \p{Canonical_Combining_Class: 10} \p{Canonical_Combining_Class=
                CCC10} (1)

\p{Canonical_Combining_Class: CCC10} (Short: \p{Ccc=CCC10}) (1:
U+05B0)

T \p{Canonical_Combining_Class: 11} \p{Canonical_Combining_Class=
CCC11} (1)

\p{Canonical_Combining_Class: CCC11} (Short: \p{Ccc=CCC11}) (1:
U+05B1)

T \p{Canonical_Combining_Class: 12} \p{Canonical_Combining_Class=
CCC12} (1)

\p{Canonical_Combining_Class: CCC12} (Short: \p{Ccc=CCC12}) (1:
U+05B2)

T \p{Canonical_Combining_Class: 13} \p{Canonical_Combining_Class=
CCC13} (1)

\p{Canonical_Combining_Class: CCC13} (Short: \p{Ccc=CCC13}) (1:
U+05B3)

T \p{Canonical_Combining_Class: 14} \p{Canonical_Combining_Class=
CCC14} (1)

\p{Canonical_Combining_Class: CCC14} (Short: \p{Ccc=CCC14}) (1:
U+05B4)

T \p{Canonical_Combining_Class: 15} \p{Canonical_Combining_Class=
CCC15} (1)

\p{Canonical_Combining_Class: CCC15} (Short: \p{Ccc=CCC15}) (1:
U+05B5)

T \p{Canonical_Combining_Class: 16} \p{Canonical_Combining_Class=
CCC16} (1)

\p{Canonical_Combining_Class: CCC16} (Short: \p{Ccc=CCC16}) (1:
U+05B6)

T \p{Canonical_Combining_Class: 17} \p{Canonical_Combining_Class=
CCC17} (1)

\p{Canonical_Combining_Class: CCC17} (Short: \p{Ccc=CCC17}) (1:
U+05B7)

T \p{Canonical_Combining_Class: 18} \p{Canonical_Combining_Class=
CCC18} (2)

\p{Canonical_Combining_Class: CCC18} (Short: \p{Ccc=CCC18}) (2:
U+05B8, U+05C7)

T \p{Canonical_Combining_Class: 19} \p{Canonical_Combining_Class=
CCC19} (2)

\p{Canonical_Combining_Class: CCC19} (Short: \p{Ccc=CCC19}) (2:
U+05B9..05BA)

T \p{Canonical_Combining_Class: 20} \p{Canonical_Combining_Class=
CCC20} (1)

    \p{Canonical_Combining_Class: CCC20} (Short: \p{Ccc=CCC20}) (1:
                U+05BB)
  T \p{Canonical_Combining_Class: 21} \p{Canonical_Combining_Class=
                CCC21} (1)
    \p{Canonical_Combining_Class: CCC21} (Short: \p{Ccc=CCC21}) (1:
                U+05BC)
  T \p{Canonical_Combining_Class: 22} \p{Canonical_Combining_Class=
                CCC22} (1)
    \p{Canonical_Combining_Class: CCC22} (Short: \p{Ccc=CCC22}) (1:
                U+05BD)
  T \p{Canonical_Combining_Class: 23} \p{Canonical_Combining_Class=
                CCC23} (1)
    \p{Canonical_Combining_Class: CCC23} (Short: \p{Ccc=CCC23}) (1:
                U+05BF)
  T \p{Canonical_Combining_Class: 24} \p{Canonical_Combining_Class=
                CCC24} (1)
    \p{Canonical_Combining_Class: CCC24} (Short: \p{Ccc=CCC24}) (1:
                U+05C1)
  T \p{Canonical_Combining_Class: 25} \p{Canonical_Combining_Class=
                CCC25} (1)
    \p{Canonical_Combining_Class: CCC25} (Short: \p{Ccc=CCC25}) (1:
                U+05C2)
  T \p{Canonical_Combining_Class: 26} \p{Canonical_Combining_Class=
                CCC26} (1)
    \p{Canonical_Combining_Class: CCC26} (Short: \p{Ccc=CCC26}) (1:
                U+FB1E)
  T \p{Canonical_Combining_Class: 27} \p{Canonical_Combining_Class=
                CCC27} (2)
    \p{Canonical_Combining_Class: CCC27} (Short: \p{Ccc=CCC27}) (2:
                U+064B, U+08F0)
  T \p{Canonical_Combining_Class: 28} \p{Canonical_Combining_Class=
                CCC28} (2)
    \p{Canonical_Combining_Class: CCC28} (Short: \p{Ccc=CCC28}) (2:
                U+064C, U+08F1)
  T \p{Canonical_Combining_Class: 29} \p{Canonical_Combining_Class=
                CCC29} (2)
    \p{Canonical_Combining_Class: CCC29} (Short: \p{Ccc=CCC29}) (2:
                U+064D, U+08F2)
  T \p{Canonical_Combining_Class: 30} \p{Canonical_Combining_Class=
                CCC30} (2)

\p{Canonical_Combining_Class: CCC30} (Short: \p{Ccc=CCC30}) (2:
              U+0618, U+064E)
T \p{Canonical_Combining_Class: 31} \p{Canonical_Combining_Class=
              CCC31} (2)
\p{Canonical_Combining_Class: CCC31} (Short: \p{Ccc=CCC31}) (2:
              U+0619, U+064F)
T \p{Canonical_Combining_Class: 32} \p{Canonical_Combining_Class=
              CCC32} (2)
\p{Canonical_Combining_Class: CCC32} (Short: \p{Ccc=CCC32}) (2:
              U+061A, U+0650)
T \p{Canonical_Combining_Class: 33} \p{Canonical_Combining_Class=
              CCC33} (1)
\p{Canonical_Combining_Class: CCC33} (Short: \p{Ccc=CCC33}) (1:
              U+0651)
T \p{Canonical_Combining_Class: 34} \p{Canonical_Combining_Class=
              CCC34} (1)
\p{Canonical_Combining_Class: CCC34} (Short: \p{Ccc=CCC34}) (1:
              U+0652)
T \p{Canonical_Combining_Class: 35} \p{Canonical_Combining_Class=
              CCC35} (1)
\p{Canonical_Combining_Class: CCC35} (Short: \p{Ccc=CCC35}) (1:
              U+0670)
T \p{Canonical_Combining_Class: 36} \p{Canonical_Combining_Class=
              CCC36} (1)
\p{Canonical_Combining_Class: CCC36} (Short: \p{Ccc=CCC36}) (1:
              U+0711)
T \p{Canonical_Combining_Class: 84} \p{Canonical_Combining_Class=
              CCC84} (1)
\p{Canonical_Combining_Class: CCC84} (Short: \p{Ccc=CCC84}) (1:
              U+0C55)
T \p{Canonical_Combining_Class: 91} \p{Canonical_Combining_Class=
              CCC91} (1)
\p{Canonical_Combining_Class: CCC91} (Short: \p{Ccc=CCC91}) (1:
              U+0C56)
T \p{Canonical_Combining_Class: 103} \p{Canonical_Combining_Class=
              CCC103} (2)
\p{Canonical_Combining_Class: CCC103} (Short: \p{Ccc=CCC103}) (2:
              U+0E38..0E39)
T \p{Canonical_Combining_Class: 107} \p{Canonical_Combining_Class=
              CCC107} (4)

   \p{Canonical_Combining_Class: CCC107} (Short: \p{Ccc=CCC107}) (4:
                U+0E48..0E4B)
 T \p{Canonical_Combining_Class: 118} \p{Canonical_Combining_Class=
                CCC118} (2)
   \p{Canonical_Combining_Class: CCC118} (Short: \p{Ccc=CCC118}) (2:
                U+0EB8..0EB9)
 T \p{Canonical_Combining_Class: 122} \p{Canonical_Combining_Class=
                CCC122} (4)
   \p{Canonical_Combining_Class: CCC122} (Short: \p{Ccc=CCC122}) (4:
                U+0EC8..0ECB)
 T \p{Canonical_Combining_Class: 129} \p{Canonical_Combining_Class=
                CCC129} (1)
   \p{Canonical_Combining_Class: CCC129} (Short: \p{Ccc=CCC129}) (1:
                U+0F71)
 T \p{Canonical_Combining_Class: 130} \p{Canonical_Combining_Class=
                CCC130} (6)
   \p{Canonical_Combining_Class: CCC130} (Short: \p{Ccc=CCC130}) (6:
                U+0F72, U+0F7A..0F7D, U+0F80)
 T \p{Canonical_Combining_Class: 132} \p{Canonical_Combining_Class=
                CCC132} (1)
   \p{Canonical_Combining_Class: CCC132} (Short: \p{Ccc=CCC132}) (1:
                U+0F74)
 T \p{Canonical_Combining_Class: 133} \p{Canonical_Combining_Class=
                CCC133} (0)
   \p{Canonical_Combining_Class: CCC133} (Short: \p{Ccc=CCC133}) (0)
 T \p{Canonical_Combining_Class: 200} \p{Canonical_Combining_Class=
                Attached_Below_Left} (0)
 T \p{Canonical_Combining_Class: 202} \p{Canonical_Combining_Class=
                Attached_Below} (5)
 T \p{Canonical_Combining_Class: 214} \p{Canonical_Combining_Class=
                Attached_Above} (1)
 T \p{Canonical_Combining_Class: 216} \p{Canonical_Combining_Class=
                Attached_Above_Right} (9)
 T \p{Canonical_Combining_Class: 218} \p{Canonical_Combining_Class=
                Below_Left} (1)
 T \p{Canonical_Combining_Class: 220} \p{Canonical_Combining_Class=
                Below} (165)
 T \p{Canonical_Combining_Class: 222} \p{Canonical_Combining_Class=
                Below_Right} (4)
 T \p{Canonical_Combining_Class: 224} \p{Canonical_Combining_Class=

Left} (2)

T \p{Canonical_Combining_Class: 226} \p{Canonical_Combining_Class=
Right} (1)

T \p{Canonical_Combining_Class: 228} \p{Canonical_Combining_Class=
Above_Left} (5)

T \p{Canonical_Combining_Class: 230} \p{Canonical_Combining_Class=
Above} (484)

T \p{Canonical_Combining_Class: 232} \p{Canonical_Combining_Class=
Above_Right} (5)

T \p{Canonical_Combining_Class: 233} \p{Canonical_Combining_Class=
Double_Below} (4)

T \p{Canonical_Combining_Class: 234} \p{Canonical_Combining_Class=
Double_Above} (5)

T \p{Canonical_Combining_Class: 240} \p{Canonical_Combining_Class=
Iota_Subscript} (1)

\p{Canonical_Combining_Class: A} \p{Canonical_Combining_Class=
Above} (484)

\p{Canonical_Combining_Class: Above} (Short: \p{Ccc=A}) (484:
U+0300..0314, U+033D..0344, U+0346,
U+034A..034C, U+0350..0352, U+0357 ...)

\p{Canonical_Combining_Class: Above_Left} (Short: \p{Ccc=AL}) (5:
U+05AE, U+18A9, U+1DF7..1DF8, U+302B)

\p{Canonical_Combining_Class: Above_Right} (Short: \p{Ccc=AR}) (5:
U+0315, U+031A, U+0358, U+1DF6, U+302C)

\p{Canonical_Combining_Class: AL} \p{Canonical_Combining_Class=
Above_Left} (5)

\p{Canonical_Combining_Class: AR} \p{Canonical_Combining_Class=
Above_Right} (5)

\p{Canonical_Combining_Class: ATA} \p{Canonical_Combining_Class=
Attached_Above} (1)

\p{Canonical_Combining_Class: ATAR} \p{Canonical_Combining_Class=
Attached_Above_Right} (9)

\p{Canonical_Combining_Class: ATB} \p{Canonical_Combining_Class=
Attached_Below} (5)

\p{Canonical_Combining_Class: ATBL} \p{Canonical_Combining_Class=
Attached_Below_Left} (0)

\p{Canonical_Combining_Class: Attached_Above} (Short: \p{Ccc=ATA})
(1: U+1DCE)

\p{Canonical_Combining_Class: Attached_Above_Right} (Short:
\p{Ccc=ATAR}) (9: U+031B, U+0F39,

U+1D165..1D166, U+1D16E..1D172)
\p{Canonical_Combining_Class: Attached_Below} (Short: \p{Ccc=ATB})
              (5: U+0321..0322, U+0327..0328, U+1DD0)
\p{Canonical_Combining_Class: Attached_Below_Left} (Short: \p{Ccc=
              ATBL}) (0)
\p{Canonical_Combining_Class: B} \p{Canonical_Combining_Class=
              Below} (165)
\p{Canonical_Combining_Class: Below} (Short: \p{Ccc=B}) (165:
              U+0316..0319, U+031C..0320,
              U+0323..0326, U+0329..0333,
              U+0339..033C, U+0347..0349 ...)
\p{Canonical_Combining_Class: Below_Left} (Short: \p{Ccc=BL}) (1:
              U+302A)
\p{Canonical_Combining_Class: Below_Right} (Short: \p{Ccc=BR}) (4:
              U+059A, U+05AD, U+1939, U+302D)
\p{Canonical_Combining_Class: BL} \p{Canonical_Combining_Class=
              Below_Left} (1)
\p{Canonical_Combining_Class: BR} \p{Canonical_Combining_Class=
              Below_Right} (4)
\p{Canonical_Combining_Class: DA} \p{Canonical_Combining_Class=
              Double_Above} (5)
\p{Canonical_Combining_Class: DB} \p{Canonical_Combining_Class=
              Double_Below} (4)
\p{Canonical_Combining_Class: Double_Above} (Short: \p{Ccc=DA})
              (5: U+035D..035E, U+0360..0361, U+1DCD)
\p{Canonical_Combining_Class: Double_Below} (Short: \p{Ccc=DB})
              (4: U+035C, U+035F, U+0362, U+1DFC)
\p{Canonical_Combining_Class: Han_Reading} (Short: \p{Ccc=HANR})
              (2: U+16FF0..16FF1)
\p{Canonical_Combining_Class: HANR} \p{Canonical_Combining_Class=
              Han_Reading} (2)
\p{Canonical_Combining_Class: Iota_Subscript} (Short: \p{Ccc=IS})
              (1: U+0345)
\p{Canonical_Combining_Class: IS} \p{Canonical_Combining_Class=
              Iota_Subscript} (1)
\p{Canonical_Combining_Class: Kana_Voicing} (Short: \p{Ccc=KV})
              (2: U+3099..309A)
\p{Canonical_Combining_Class: KV} \p{Canonical_Combining_Class=
              Kana_Voicing} (2)
\p{Canonical_Combining_Class: L} \p{Canonical_Combining_Class=

Left} (2)

\p{Canonical_Combining_Class: Left} (Short: \p{Ccc=L}) (2:
           U+302E..302F)
\p{Canonical_Combining_Class: NK} \p{Canonical_Combining_Class=
           Nukta} (26)
\p{Canonical_Combining_Class: Not_Reordered} (Short: \p{Ccc=NR})
           (1_113_240 plus all above-Unicode code
           points: U+0000..02FF, U+034F,
           U+0370..0482, U+0488..0590, U+05BE,
           U+05C0 ...)
\p{Canonical_Combining_Class: NR} \p{Canonical_Combining_Class=
           Not_Reordered} (1_113_240 plus all
           above-Unicode code points)
\p{Canonical_Combining_Class: Nukta} (Short: \p{Ccc=NK}) (26:
           U+093C, U+09BC, U+0A3C, U+0ABC, U+0B3C,
           U+0CBC ...)
\p{Canonical_Combining_Class: OV} \p{Canonical_Combining_Class=
           Overlay} (32)
\p{Canonical_Combining_Class: Overlay} (Short: \p{Ccc=OV}) (32:
           U+0334..0338, U+1CD4, U+1CE2..1CE8,
           U+20D2..20D3, U+20D8..20DA, U+20E5..20E6
           ...)
\p{Canonical_Combining_Class: R} \p{Canonical_Combining_Class=
           Right} (1)
\p{Canonical_Combining_Class: Right} (Short: \p{Ccc=R}) (1:
           U+1D16D)
\p{Canonical_Combining_Class: Virama} (Short: \p{Ccc=VR}) (61:
           U+094D, U+09CD, U+0A4D, U+0ACD, U+0B4D,
           U+0BCD ...)
\p{Canonical_Combining_Class: VR} \p{Canonical_Combining_Class=
           Virama} (61)
\p{Cans}   \p{Canadian_Aboriginal} (=
           \p{Script_Extensions=
           Canadian_Aboriginal}) (710)
\p{Cari}   \p{Carian} (= \p{Script_Extensions=
           Carian}) (NOT \p{Block=Carian}) (49)
\p{Carian}   \p{Script_Extensions=Carian} (Short:
           \p{Cari}; NOT \p{Block=Carian}) (49)
\p{Case_Ignorable}   \p{Case_Ignorable=Y} (Short: \p{CI}) (2413)
\p{Case_Ignorable: N*} (Short: \p{CI=N}, \P{CI}) (1_111_699 plus

all above-Unicode code points: [\x00-
\x20!\"#\$\%&\(\)*+,\-\/0-9;<=>?\@A-Z
\[\\\]_a-z\{\|\}~\x7f-\xa7\xa9-\xac\xae
\xb0-\xb3\xb5-\xb6\xb9-\xff],
U+0100..02AF, U+0370..0373,
U+0376..0379, U+037B..0383, U+0386 ...)

\p{Case_Ignorable: Y*} (Short: \p{CI=Y}, \p{CI}) (2413: [\'.:\^`
\xa8\xad\xaf\xb4\xb7-\xb8],
U+02B0..036F, U+0374..0375, U+037A,
U+0384..0385, U+0387 ...)

\p{Cased}   \p{Cased=Y} (4286)

\p{Cased: N*}   (Single: \P{Cased}) (1_109_826 plus all
above-Unicode code points: [\x00-\x20!
\"#\$\%&\'\(\)*+,\-.\/0-9:;<=>?\@\[\\\]
\^_`\{\|\}~\x7f-\xa9\xab-\xb4\xb6-\xb9
\xbb-\xbf\xd7\xf7], U+01BB,
U+01C0..01C3, U+0294, U+02B9..02BF,
U+02C2..02DF ...)

\p{Cased: Y*}   (Single: \p{Cased}) (4286: [A-Za-z\xaa
\xb5\xba\xc0-\xd6\xd8-\xf6\xf8-\xff],
U+0100..01BA, U+01BC..01BF,
U+01C4..0293, U+0295..02B8, U+02C0..02C1
...)

\p{Cased_Letter}   \p{General_Category=Cased_Letter} (Short:
\p{LC}) (3977)

\p{Category: *}   \p{General_Category: *}

\p{Caucasian_Albanian} \p{Script_Extensions=Caucasian_Albanian}
(Short: \p{Aghb}; NOT \p{Block=
Caucasian_Albanian}) (53)

\p{Cc}   \p{XPosixCntrl} (= \p{General_Category=
Control}) (65)

\p{Ccc: *}   \p{Canonical_Combining_Class: *}

\p{CE}   \p{Composition_Exclusion} (=
\p{Composition_Exclusion=Y}) (81)

\p{CE: *}   \p{Composition_Exclusion: *}

\p{Cf}   \p{Format} (= \p{General_Category=Format})
(161)

\p{Chakma}   \p{Script_Extensions=Chakma} (Short:
\p{Cakm}; NOT \p{Block=Chakma}) (91)

\p{Cham}   \p{Script_Extensions=Cham} (NOT \p{Block=

                        Cham}) (83)
\p{Changes_When_Casefolded} \p{Changes_When_Casefolded=Y} (Short:
                \p{CWCF}) (1466)
\p{Changes_When_Casefolded: N*} (Short: \p{CWCF=N}, \P{CWCF})
                (1_112_646 plus all above-Unicode code
                points: [\x00-\x20!\"#\$\%&\'\(\)*+,\-.
                \/0-9:;<=>?\@\[\\\]\^_`a-z\{\|\}~\x7f-
                \xb4\xb6-\xbf\xd7\xe0-\xff], U+0101,
                U+0103, U+0105, U+0107, U+0109 ...)
\p{Changes_When_Casefolded: Y*} (Short: \p{CWCF=Y}, \p{CWCF})
                (1466: [A-Z\xb5\xc0-\xd6\xd8-\xdf],
                U+0100, U+0102, U+0104, U+0106, U+0108
                ...)
\p{Changes_When_Casemapped} \p{Changes_When_Casemapped=Y} (Short:
                \p{CWCM}) (2847)
\p{Changes_When_Casemapped: N*} (Short: \p{CWCM=N}, \P{CWCM})
                (1_111_265 plus all above-Unicode code
                points: [\x00-\x20!\"#\$\%&\'\(\)*+,\-.
                \/0-9:;<=>?\@\[\\\]\^_`\{\|\}~\x7f-\xb4
                \xb6-\xbf\xd7\xf7], U+0138, U+018D,
                U+019B, U+01AA..01AB, U+01BA..01BB ...)
\p{Changes_When_Casemapped: Y*} (Short: \p{CWCM=Y}, \p{CWCM})
                (2847: [A-Za-z\xb5\xc0-\xd6\xd8-\xf6
                \xf8-\xff], U+0100..0137, U+0139..018C,
                U+018E..019A, U+019C..01A9, U+01AC..01B9
                ...)
\p{Changes_When_Lowercased} \p{Changes_When_Lowercased=Y} (Short:
                \p{CWL}) (1393)
\p{Changes_When_Lowercased: N*} (Short: \p{CWL=N}, \P{CWL})
                (1_112_719 plus all above-Unicode code
                points: [\x00-\x20!\"#\$\%&\'\(\)*+,\-.
                \/0-9:;<=>?\@\[\\\]\^_`a-z\{\|\}~\x7f-
                \xbf\xd7\xdf-\xff], U+0101, U+0103,
                U+0105, U+0107, U+0109 ...)
\p{Changes_When_Lowercased: Y*} (Short: \p{CWL=Y}, \p{CWL}) (1393:
                [A-Z\xc0-\xd6\xd8-\xde], U+0100, U+0102,
                U+0104, U+0106, U+0108 ...)
\p{Changes_When_NFKC_Casefolded} \p{Changes_When_NFKC_Casefolded=
                Y} (Short: \p{CWKCF}) (10_329)
\p{Changes_When_NFKC_Casefolded: N*} (Short: \p{CWKCF=N},

          \P{CWKCF}) (1_103_783 plus all above-
          Unicode code points: [\x00-\x20!\"#\$
          \%&\'\(\)*+,\-.\/0-9:;<=>?\@\[\\\]\^_'a-
          z\{\|\}~\x7f-\x9f\xa1-\xa7\xa9\xab-\xac
          \xae\xb0-\xb1\xb6-\xb7\xbb\xbf\xd7\xe0-
          \xff], U+0101, U+0103, U+0105, U+0107,
          U+0109 ...)

\p{Changes_When_NFKC_Casefolded: Y*} (Short: \p{CWKCF=Y},
          \p{CWKCF}) (10_329: [A-Z\xa0\xa8\xaa
          \xad\xaf\xb2-\xb5\xb8-\xba\xbc-\xbe\xc0-
          \xd6\xd8-\xdf], U+0100, U+0102, U+0104,
          U+0106, U+0108 ...)

\p{Changes_When_Titlecased} \p{Changes_When_Titlecased=Y} (Short:
          \p{CWT}) (1412)

\p{Changes_When_Titlecased: N*} (Short: \p{CWT=N}, \P{CWT})
          (1_112_700 plus all above-Unicode code
          points: [\x00-\x20!\"#\$\%&\'\(\)*+,\-.
          \/0-9:;<=>?\@A-Z\[\\\]\^_'\{\|\}~\x7f-
          \xb4\xb6-\xde\xf7], U+0100, U+0102,
          U+0104, U+0106, U+0108 ...)

\p{Changes_When_Titlecased: Y*} (Short: \p{CWT=Y}, \p{CWT}) (1412:
          [a-z\xb5\xdf-\xf6\xf8-\xff], U+0101,
          U+0103, U+0105, U+0107, U+0109 ...)

\p{Changes_When_Uppercased} \p{Changes_When_Uppercased=Y} (Short:
          \p{CWU}) (1485)

\p{Changes_When_Uppercased: N*} (Short: \p{CWU=N}, \P{CWU})
          (1_112_627 plus all above-Unicode code
          points: [\x00-\x20!\"#\$\%&\'\(\)*+,\-.
          \/0-9:;<=>?\@A-Z\[\\\]\^_'\{\|\}~\x7f-
          \xb4\xb6-\xde\xf7], U+0100, U+0102,
          U+0104, U+0106, U+0108 ...)

\p{Changes_When_Uppercased: Y*} (Short: \p{CWU=Y}, \p{CWU}) (1485:
          [a-z\xb5\xdf-\xf6\xf8-\xff], U+0101,
          U+0103, U+0105, U+0107, U+0109 ...)

\p{Cher}   \p{Cherokee} (= \p{Script_Extensions=
          Cherokee}) (NOT \p{Block=Cherokee}) (172)

\p{Cherokee}   \p{Script_Extensions=Cherokee} (Short:
          \p{Cher}; NOT \p{Block=Cherokee}) (172)

X \p{Cherokee_Sup}   \p{Cherokee_Supplement} (= \p{Block=
          Cherokee_Supplement}) (80)

X \p{Cherokee_Supplement} \p{Block=Cherokee_Supplement} (Short:
            \p{InCherokeeSup}) (80)
X \p{Chess_Symbols}   \p{Block=Chess_Symbols} (112)
 \p{Chorasmian}   \p{Script_Extensions=Chorasmian} (Short:
            \p{Chrs}; NOT \p{Block=Chorasmian}) (28)
 \p{Chrs}   \p{Chorasmian} (= \p{Script_Extensions=
            Chorasmian}) (NOT \p{Block=Chorasmian})
            (28)
 \p{CI}   \p{Case_Ignorable} (= \p{Case_Ignorable=
            Y}) (2413)
 \p{CI: *}   \p{Case_Ignorable: *}
X \p{CJK}   \p{CJK_Unified_Ideographs} (= \p{Block=
            CJK_Unified_Ideographs}) (20_992)
X \p{CJK_Compat}   \p{CJK_Compatibility} (= \p{Block=
            CJK_Compatibility}) (256)
X \p{CJK_Compat_Forms}   \p{CJK_Compatibility_Forms} (= \p{Block=
            CJK_Compatibility_Forms}) (32)
X \p{CJK_Compat_Ideographs} \p{CJK_Compatibility_Ideographs} (=
            \p{Block=CJK_Compatibility_Ideographs})
            (512)
X \p{CJK_Compat_Ideographs_Sup}
            \p{CJK_Compatibility_Ideographs_-
            Supplement} (= \p{Block=
            CJK_Compatibility_Ideographs_-
            Supplement}) (544)
X \p{CJK_Compatibility} \p{Block=CJK_Compatibility} (Short:
            \p{InCJKCompat}) (256)
X \p{CJK_Compatibility_Forms} \p{Block=CJK_Compatibility_Forms}
            (Short: \p{InCJKCompatForms}) (32)
X \p{CJK_Compatibility_Ideographs} \p{Block=
            CJK_Compatibility_Ideographs} (Short:
            \p{InCJKCompatIdeographs}) (512)
X \p{CJK_Compatibility_Ideographs_Supplement} \p{Block=
            CJK_Compatibility_Ideographs_Supplement}
            (Short: \p{InCJKCompatIdeographsSup})
            (544)
X \p{CJK_Ext_A}   \p{CJK_Unified_Ideographs_Extension_A} (=
            \p{Block=
            CJK_Unified_Ideographs_Extension_A})
            (6592)

```
X \p{CJK_Ext_B}        \p{CJK_Unified_Ideographs_Extension_B} (=
                \p{Block=
                CJK_Unified_Ideographs_Extension_B})
                (42_720)
X \p{CJK_Ext_C}        \p{CJK_Unified_Ideographs_Extension_C} (=
                \p{Block=
                CJK_Unified_Ideographs_Extension_C})
                (4160)
X \p{CJK_Ext_D}        \p{CJK_Unified_Ideographs_Extension_D} (=
                \p{Block=
                CJK_Unified_Ideographs_Extension_D})
                (224)
X \p{CJK_Ext_E}        \p{CJK_Unified_Ideographs_Extension_E} (=
                \p{Block=
                CJK_Unified_Ideographs_Extension_E})
                (5776)
X \p{CJK_Ext_F}        \p{CJK_Unified_Ideographs_Extension_F} (=
                \p{Block=
                CJK_Unified_Ideographs_Extension_F})
                (7488)
X \p{CJK_Ext_G}        \p{CJK_Unified_Ideographs_Extension_G} (=
                \p{Block=
                CJK_Unified_Ideographs_Extension_G})
                (4944)
X \p{CJK_Radicals_Sup}   \p{CJK_Radicals_Supplement} (= \p{Block=
                CJK_Radicals_Supplement}) (128)
X \p{CJK_Radicals_Supplement} \p{Block=CJK_Radicals_Supplement}
                (Short: \p{InCJKRadicalsSup}) (128)
X \p{CJK_Strokes}       \p{Block=CJK_Strokes} (48)
X \p{CJK_Symbols}       \p{CJK_Symbols_And_Punctuation} (=
                \p{Block=CJK_Symbols_And_Punctuation})
                (64)
X \p{CJK_Symbols_And_Punctuation} \p{Block=
                CJK_Symbols_And_Punctuation} (Short:
                \p{InCJKSymbols}) (64)
X \p{CJK_Unified_Ideographs} \p{Block=CJK_Unified_Ideographs}
                (Short: \p{InCJK}) (20_992)
X \p{CJK_Unified_Ideographs_Extension_A} \p{Block=
                CJK_Unified_Ideographs_Extension_A}
                (Short: \p{InCJKExtA}) (6592)
```

X \p{CJK_Unified_Ideographs_Extension_B} \p{Block=
             CJK_Unified_Ideographs_Extension_B}
             (Short: \p{InCJKExtB}) (42_720)
X \p{CJK_Unified_Ideographs_Extension_C} \p{Block=
             CJK_Unified_Ideographs_Extension_C}
             (Short: \p{InCJKExtC}) (4160)
X \p{CJK_Unified_Ideographs_Extension_D} \p{Block=
             CJK_Unified_Ideographs_Extension_D}
             (Short: \p{InCJKExtD}) (224)
X \p{CJK_Unified_Ideographs_Extension_E} \p{Block=
             CJK_Unified_Ideographs_Extension_E}
             (Short: \p{InCJKExtE}) (5776)
X \p{CJK_Unified_Ideographs_Extension_F} \p{Block=
             CJK_Unified_Ideographs_Extension_F}
             (Short: \p{InCJKExtF}) (7488)
X \p{CJK_Unified_Ideographs_Extension_G} \p{Block=
             CJK_Unified_Ideographs_Extension_G}
             (Short: \p{InCJKExtG}) (4944)
 \p{Close_Punctuation}   \p{General_Category=Close_Punctuation}
             (Short: \p{Pe}) (73)
 \p{Cn}   \p{Unassigned} (= \p{General_Category=
             Unassigned}) (830_672 plus all above-
             Unicode code points)
 \p{Cntrl}   \p{XPosixCntrl} (= \p{General_Category=
             Control}) (65)
 \p{Co}   \p{Private_Use} (= \p{General_Category=
             Private_Use}) (NOT \p{Private_Use_Area})
             (137_468)
X \p{Combining_Diacritical_Marks} \p{Block=
             Combining_Diacritical_Marks} (Short:
             \p{InDiacriticals}) (112)
X \p{Combining_Diacritical_Marks_Extended} \p{Block=
             Combining_Diacritical_Marks_Extended}
             (Short: \p{InDiacriticalsExt}) (80)
X \p{Combining_Diacritical_Marks_For_Symbols} \p{Block=
             Combining_Diacritical_Marks_For_Symbols}
             (Short: \p{InDiacriticalsForSymbols})
             (48)
X \p{Combining_Diacritical_Marks_Supplement} \p{Block=
             Combining_Diacritical_Marks_Supplement}

                    (Short: \p{InDiacriticalsSup}) (64)
   X \p{Combining_Half_Marks} \p{Block=Combining_Half_Marks} (Short:
                    \p{InHalfMarks}) (16)
    \p{Combining_Mark}   \p{Mark} (= \p{General_Category=Mark})
                    (2295)
   X \p{Combining_Marks_For_Symbols}
                    \p{Combining_Diacritical_Marks_For_-
                    Symbols} (= \p{Block=
                    Combining_Diacritical_Marks_For_-
                    Symbols}) (48)
    \p{Common}   \p{Script_Extensions=Common} (Short:
                    \p{Zyyy}) (7661)
   X \p{Common_Indic_Number_Forms} \p{Block=Common_Indic_Number_Forms}
                    (Short: \p{InIndicNumberForms}) (16)
    \p{Comp_Ex}   \p{Full_Composition_Exclusion} (=
                    \p{Full_Composition_Exclusion=Y}) (1120)
    \p{Comp_Ex: *}   \p{Full_Composition_Exclusion: *}
   X \p{Compat_Jamo}   \p{Hangul_Compatibility_Jamo} (= \p{Block=
                    Hangul_Compatibility_Jamo}) (96)
    \p{Composition_Exclusion} \p{Composition_Exclusion=Y} (Short:
                    \p{CE}) (81)
    \p{Composition_Exclusion: N*} (Short: \p{CE=N}, \P{CE}) (1_114_031
                    plus all above-Unicode code points:
                    U+0000..0957, U+0960..09DB, U+09DE,
                    U+09E0..0A32, U+0A34..0A35, U+0A37..0A58
                    ...)
    \p{Composition_Exclusion: Y*} (Short: \p{CE=Y}, \p{CE}) (81:
                    U+0958..095F, U+09DC..09DD, U+09DF,
                    U+0A33, U+0A36, U+0A59..0A5B ...)
    \p{Connector_Punctuation} \p{General_Category=
                    Connector_Punctuation} (Short: \p{Pc})
                    (10)
    \p{Control}   \p{XPosixCntrl} (= \p{General_Category=
                    Control}) (65)
   X \p{Control_Pictures}   \p{Block=Control_Pictures} (64)
    \p{Copt}   \p{Coptic} (= \p{Script_Extensions=
                    Coptic}) (NOT \p{Block=Coptic}) (165)
    \p{Coptic}   \p{Script_Extensions=Coptic} (Short:
                    \p{Copt}; NOT \p{Block=Coptic}) (165)
   X \p{Coptic_Epact_Numbers} \p{Block=Coptic_Epact_Numbers} (32)

X \p{Counting_Rod}   \p{Counting_Rod_Numerals} (= \p{Block=
                Counting_Rod_Numerals}) (32)
X \p{Counting_Rod_Numerals} \p{Block=Counting_Rod_Numerals} (Short:
                \p{InCountingRod}) (32)
 \p{Cprt}   \p{Cypriot} (= \p{Script_Extensions=
                Cypriot}) (112)
 \p{Cs}   \p{Surrogate} (= \p{General_Category=
                Surrogate}) (2048)
 \p{Cuneiform}   \p{Script_Extensions=Cuneiform} (Short:
                \p{Xsux}; NOT \p{Block=Cuneiform}) (1234)
X \p{Cuneiform_Numbers}   \p{Cuneiform_Numbers_And_Punctuation} (=
                \p{Block=
                Cuneiform_Numbers_And_Punctuation}) (128)
X \p{Cuneiform_Numbers_And_Punctuation} \p{Block=
                Cuneiform_Numbers_And_Punctuation}
                (Short: \p{InCuneiformNumbers}) (128)
 \p{Currency_Symbol}   \p{General_Category=Currency_Symbol}
                (Short: \p{Sc}) (62)
X \p{Currency_Symbols}   \p{Block=Currency_Symbols} (48)
 \p{CWCF}   \p{Changes_When_Casefolded} (=
                \p{Changes_When_Casefolded=Y}) (1466)
 \p{CWCF: *}   \p{Changes_When_Casefolded: *}
 \p{CWCM}   \p{Changes_When_Casemapped} (=
                \p{Changes_When_Casemapped=Y}) (2847)
 \p{CWCM: *}   \p{Changes_When_Casemapped: *}
 \p{CWKCF}   \p{Changes_When_NFKC_Casefolded} (=
                \p{Changes_When_NFKC_Casefolded=Y})
                (10_329)
 \p{CWKCF: *}   \p{Changes_When_NFKC_Casefolded: *}
 \p{CWL}   \p{Changes_When_Lowercased} (=
                \p{Changes_When_Lowercased=Y}) (1393)
 \p{CWL: *}   \p{Changes_When_Lowercased: *}
 \p{CWT}   \p{Changes_When_Titlecased} (=
                \p{Changes_When_Titlecased=Y}) (1412)
 \p{CWT: *}   \p{Changes_When_Titlecased: *}
 \p{CWU}   \p{Changes_When_Uppercased} (=
                \p{Changes_When_Uppercased=Y}) (1485)
 \p{CWU: *}   \p{Changes_When_Uppercased: *}
 \p{Cypriot}   \p{Script_Extensions=Cypriot} (Short:
                \p{Cprt}) (112)

X \p{Cypriot_Syllabary}   \p{Block=Cypriot_Syllabary} (64)

 \p{Cyrillic}   \p{Script_Extensions=Cyrillic} (Short:
               \p{Cyrl}; NOT \p{Block=Cyrillic}) (447)

X \p{Cyrillic_Ext_A}   \p{Cyrillic_Extended_A} (= \p{Block=
               Cyrillic_Extended_A}) (32)

X \p{Cyrillic_Ext_B}   \p{Cyrillic_Extended_B} (= \p{Block=
               Cyrillic_Extended_B}) (96)

X \p{Cyrillic_Ext_C}   \p{Cyrillic_Extended_C} (= \p{Block=
               Cyrillic_Extended_C}) (16)

X \p{Cyrillic_Extended_A} \p{Block=Cyrillic_Extended_A} (Short:
               \p{InCyrillicExtA}) (32)

X \p{Cyrillic_Extended_B} \p{Block=Cyrillic_Extended_B} (Short:
               \p{InCyrillicExtB}) (96)

X \p{Cyrillic_Extended_C} \p{Block=Cyrillic_Extended_C} (Short:
               \p{InCyrillicExtC}) (16)

X \p{Cyrillic_Sup}   \p{Cyrillic_Supplement} (= \p{Block=
               Cyrillic_Supplement}) (48)

X \p{Cyrillic_Supplement} \p{Block=Cyrillic_Supplement} (Short:
               \p{InCyrillicSup}) (48)

X \p{Cyrillic_Supplementary} \p{Cyrillic_Supplement} (= \p{Block=
               Cyrillic_Supplement}) (48)

 \p{Cyrl}   \p{Cyrillic} (= \p{Script_Extensions=
               Cyrillic}) (NOT \p{Block=Cyrillic}) (447)

 \p{Dash}   \p{Dash=Y} (29)

 \p{Dash: N*}   (Single: \P{Dash}) (1_114_083 plus all
               above-Unicode code points: [\x00-\x20!
               \"#\$\%&\'\(\)*+,.\/0-9:;<=>?\@A-Z
               \[\\\]\^_'a-z\{\|\}~\x7f-\xff],
               U+0100..0589, U+058B..05BD,
               U+05BF..13FF, U+1401..1805, U+1807..200F
               ...)

 \p{Dash: Y*}   (Single: \p{Dash}) (29: [\-], U+058A,
               U+05BE, U+1400, U+1806, U+2010..2015 ...)

 \p{Dash_Punctuation}   \p{General_Category=Dash_Punctuation}
               (Short: \p{Pd}) (25)

 \p{Decimal_Number}   \p{XPosixDigit} (= \p{General_Category=
               Decimal_Number}) (650)

 \p{Decomposition_Type: Can} \p{Decomposition_Type=Canonical}
               (13_233)

 \p{Decomposition_Type: Canonical} (Short: \p{Dt=Can}) (13_233:

                      [\xc0-\xc5\xc7-\xcf\xd1-\xd6\xd9-\xdd
                      \xe0-\xe5\xe7-\xef\xf1-\xf6\xf9-\xfd
                      \xff], U+0100..010F, U+0112..0125,
                      U+0128..0130, U+0134..0137, U+0139..013E
                      ...)

\p{Decomposition_Type: Circle} (Short: \p{Dt=Enc}) (240:
                      U+2460..2473, U+24B6..24EA,
                      U+3244..3247, U+3251..327E,
                      U+3280..32BF, U+32D0..32FE ...)

\p{Decomposition_Type: Com} \p{Decomposition_Type=Compat} (720)

\p{Decomposition_Type: Compat} (Short: \p{Dt=Com}) (720: [\xa8
                      \xaf\xb4-\xb5\xb8], U+0132..0133,
                      U+013F..0140, U+0149, U+017F,
                      U+01C4..01CC ...)

\p{Decomposition_Type: Enc} \p{Decomposition_Type=Circle} (240)

\p{Decomposition_Type: Fin} \p{Decomposition_Type=Final} (240)

\p{Decomposition_Type: Final} (Short: \p{Dt=Fin}) (240: U+FB51,
                      U+FB53, U+FB57, U+FB5B, U+FB5F, U+FB63
                      ...)

\p{Decomposition_Type: Font} (Short: \p{Dt=Font}) (1194: U+2102,
                      U+210A..2113, U+2115, U+2119..211D,
                      U+2124, U+2128 ...)

\p{Decomposition_Type: Fra} \p{Decomposition_Type=Fraction} (20)

\p{Decomposition_Type: Fraction} (Short: \p{Dt=Fra}) (20: [\xbc-
                      \xbe], U+2150..215F, U+2189)

\p{Decomposition_Type: Init} \p{Decomposition_Type=Initial} (171)

\p{Decomposition_Type: Initial} (Short: \p{Dt=Init}) (171: U+FB54,
                      U+FB58, U+FB5C, U+FB60, U+FB64, U+FB68
                      ...)

\p{Decomposition_Type: Iso} \p{Decomposition_Type=Isolated} (238)

\p{Decomposition_Type: Isolated} (Short: \p{Dt=Iso}) (238: U+FB50,
                      U+FB52, U+FB56, U+FB5A, U+FB5E, U+FB62
                      ...)

\p{Decomposition_Type: Med} \p{Decomposition_Type=Medial} (82)

\p{Decomposition_Type: Medial} (Short: \p{Dt=Med}) (82: U+FB55,
                      U+FB59, U+FB5D, U+FB61, U+FB65, U+FB69
                      ...)

\p{Decomposition_Type: Nar} \p{Decomposition_Type=Narrow} (122)

\p{Decomposition_Type: Narrow} (Short: \p{Dt=Nar}) (122:
                      U+FF61..FFBE, U+FFC2..FFC7,

                U+FFCA..FFCF, U+FFD2..FFD7,
                U+FFDA..FFDC, U+FFE8..FFEE)

\p{Decomposition_Type: Nb} \p{Decomposition_Type=Nobreak} (5)

\p{Decomposition_Type: Nobreak} (Short: \p{Dt=Nb}) (5: [\xa0],
                U+0F0C, U+2007, U+2011, U+202F)

\p{Decomposition_Type: Non_Canon} \p{Decomposition_Type=
                Non_Canonical} (Perl extension) (3675)

\p{Decomposition_Type: Non_Canonical} Union of all non-canonical
                decompositions (Short: \p{Dt=NonCanon})
                (Perl extension) (3675: [\xa0\xa8\xaa
                \xaf\xb2-\xb5\xb8-\xba\xbc-\xbe],
                U+0132..0133, U+013F..0140, U+0149,
                U+017F, U+01C4..01CC ...)

\p{Decomposition_Type: None} (Short: \p{Dt=None}) (1_097_204 plus
                all above-Unicode code points: [\x00-
                \x9f\xa1-\xa7\xa9\xab-\xae\xb0-\xb1\xb6-
                \xb7\xbb\xbf\xc6\xd0\xd7-\xd8\xde-\xdf
                \xe6\xf0\xf7-\xf8\xfe], U+0110..0111,
                U+0126..0127, U+0131, U+0138,
                U+0141..0142 ...)

\p{Decomposition_Type: Small} (Short: \p{Dt=Sml}) (26:
                U+FE50..FE52, U+FE54..FE66, U+FE68..FE6B)

\p{Decomposition_Type: Sml} \p{Decomposition_Type=Small} (26)

\p{Decomposition_Type: Sqr} \p{Decomposition_Type=Square} (286)

\p{Decomposition_Type: Square} (Short: \p{Dt=Sqr}) (286: U+3250,
                U+32CC..32CF, U+32FF..3357,
                U+3371..33DF, U+33FF, U+1F130..1F14F ...)

\p{Decomposition_Type: Sub} (Short: \p{Dt=Sub}) (38: U+1D62..1D6A,
                U+2080..208E, U+2090..209C, U+2C7C)

\p{Decomposition_Type: Sup} \p{Decomposition_Type=Super} (154)

\p{Decomposition_Type: Super} (Short: \p{Dt=Sup}) (154: [\xaa\xb2-
                \xb3\xb9-\xba], U+02B0..02B8,
                U+02E0..02E4, U+10FC, U+1D2C..1D2E,
                U+1D30..1D3A ...)

\p{Decomposition_Type: Vert} \p{Decomposition_Type=Vertical} (35)

\p{Decomposition_Type: Vertical} (Short: \p{Dt=Vert}) (35: U+309F,
                U+30FF, U+FE10..FE19, U+FE30..FE44,
                U+FE47..FE48)

\p{Decomposition_Type: Wide} (Short: \p{Dt=Wide}) (104: U+3000,
                U+FF01..FF60, U+FFE0..FFE6)

\p{Default_Ignorable_Code_Point} \p{Default_Ignorable_Code_Point=
                 Y} (Short: \p{DI}) (4173)
\p{Default_Ignorable_Code_Point: N*} (Short: \p{DI=N}, \P{DI})
                 (1_109_939 plus all above-Unicode code
                 points: [\x00-\xac\xae-\xff],
                 U+0100..034E, U+0350..061B,
                 U+061D..115E, U+1161..17B3, U+17B6..180A
                 ...)
\p{Default_Ignorable_Code_Point: Y*} (Short: \p{DI=Y}, \p{DI})
                 (4173: [\xad], U+034F, U+061C,
                 U+115F..1160, U+17B4..17B5, U+180B..180E
                 ...)
\p{Dep}   \p{Deprecated} (= \p{Deprecated=Y}) (15)
\p{Dep: *}   \p{Deprecated: *}
\p{Deprecated}   \p{Deprecated=Y} (Short: \p{Dep}) (15)
\p{Deprecated: N*}   (Short: \p{Dep=N}, \P{Dep}) (1_114_097
                 plus all above-Unicode code points:
                 U+0000..0148, U+014A..0672,
                 U+0674..0F76, U+0F78, U+0F7A..17A2,
                 U+17A5..2069 ...)
\p{Deprecated: Y*}   (Short: \p{Dep=Y}, \p{Dep}) (15: U+0149,
                 U+0673, U+0F77, U+0F79, U+17A3..17A4,
                 U+206A..206F ...)
\p{Deseret}   \p{Script_Extensions=Deseret} (Short:
                 \p{Dsrt}) (80)
\p{Deva}   \p{Devanagari} (= \p{Script_Extensions=
                 Devanagari}) (NOT \p{Block=Devanagari})
                 (210)
\p{Devanagari}   \p{Script_Extensions=Devanagari} (Short:
                 \p{Deva}; NOT \p{Block=Devanagari}) (210)
X \p{Devanagari_Ext}   \p{Devanagari_Extended} (= \p{Block=
                 Devanagari_Extended}) (32)
X \p{Devanagari_Extended} \p{Block=Devanagari_Extended} (Short:
                 \p{InDevanagariExt}) (32)
\p{DI}   \p{Default_Ignorable_Code_Point} (=
                 \p{Default_Ignorable_Code_Point=Y})
                 (4173)
\p{DI: *}   \p{Default_Ignorable_Code_Point: *}
\p{Dia}   \p{Diacritic} (= \p{Diacritic=Y}) (882)
\p{Dia: *}   \p{Diacritic: *}

    \p{Diacritic}   \p{Diacritic=Y} (Short: \p{Dia}) (882)
    \p{Diacritic: N*}   (Short: \p{Dia=N}, \P{Dia}) (1_113_230
                 plus all above-Unicode code points:
                 [\x00-\x20!\"#\$\%&\'\’\(\)*+,\-.\/0-9:;<=
                 >?\@A-Z\[\\\]_a-z\{\|\}~\x7f-\xa7\xa9-
                 \xae\xb0-\xb3\xb5-\xb6\xb9-\xff],
                 U+0100..02AF, U+034F, U+0358..035C,
                 U+0363..0373, U+0376..0379 ...)
     \p{Diacritic: Y*}   (Short: \p{Dia=Y}, \p{Dia}) (882: [\^‘
                 \xa8\xaf\xb4\xb7-\xb8], U+02B0..034E,
                 U+0350..0357, U+035D..0362,
                 U+0374..0375, U+037A ...)
    X \p{Diacriticals}   \p{Combining_Diacritical_Marks} (=
                 \p{Block=Combining_Diacritical_Marks})
                 (112)
    X \p{Diacriticals_Ext}   \p{Combining_Diacritical_Marks_Extended}
                 (= \p{Block=
                 Combining_Diacritical_Marks_Extended})
                 (80)
    X \p{Diacriticals_For_Symbols}
                 \p{Combining_Diacritical_Marks_For_-
                 Symbols} (= \p{Block=
                 Combining_Diacritical_Marks_For_-
                 Symbols}) (48)
    X \p{Diacriticals_Sup}   \p{Combining_Diacritical_Marks_Supplement}
                 (= \p{Block=
                 Combining_Diacritical_Marks_Supplement})
                 (64)
     \p{Diak}   \p{Dives_Akuru} (= \p{Script_Extensions=
                 Dives_Akuru}) (NOT \p{Block=
                 Dives_Akuru}) (72)
     \p{Digit}   \p{XPosixDigit} (= \p{General_Category=
                 Decimal_Number}) (650)
    X \p{Dingbats}   \p{Block=Dingbats} (192)
     \p{Dives_Akuru}   \p{Script_Extensions=Dives_Akuru} (Short:
                 \p{Diak}; NOT \p{Block=Dives_Akuru}) (72)
     \p{Dogr}   \p{Dogra} (= \p{Script_Extensions=Dogra})
                 (NOT \p{Block=Dogra}) (82)
     \p{Dogra}   \p{Script_Extensions=Dogra} (Short:
                 \p{Dogr}; NOT \p{Block=Dogra}) (82)

```
    X \p{Domino}              \p{Domino_Tiles} (= \p{Block=
                          Domino_Tiles}) (112)
    X \p{Domino_Tiles}       \p{Block=Domino_Tiles} (Short:
                          \p{InDomino}) (112)
     \p{Dsrt}                \p{Deseret} (= \p{Script_Extensions=
                          Deseret}) (80)
     \p{Dt: *}               \p{Decomposition_Type: *}
     \p{Dupl}                 \p{Duployan} (= \p{Script_Extensions=
                          Duployan}) (NOT \p{Block=Duployan}) (147)
     \p{Duployan}            \p{Script_Extensions=Duployan} (Short:
                          \p{Dupl}; NOT \p{Block=Duployan}) (147)
     \p{Ea: *}               \p{East_Asian_Width: *}
    X \p{Early_Dynastic_Cuneiform} \p{Block=Early_Dynastic_Cuneiform}
                          (208)
     \p{East_Asian_Width: A} \p{East_Asian_Width=Ambiguous} (138_739)
     \p{East_Asian_Width: Ambiguous} (Short: \p{Ea=A}) (138_739: [\xa1
                          \xa4\xa7-\xa8\xaa\xad-\xae\xb0-\xb4\xb6-
                          \xba\xbc-\xbf\xc6\xd0\xd7-\xd8\xde-\xe1
                          \xe6\xe8-\xea\xec-\xed\xf0\xf2-\xf3\xf7-
                          \xfa\xfc\xfe], U+0101, U+0111, U+0113,
                          U+011B, U+0126..0127 ...)
     \p{East_Asian_Width: F} \p{East_Asian_Width=Fullwidth} (104)
     \p{East_Asian_Width: Fullwidth} (Short: \p{Ea=F}) (104: U+3000,
                          U+FF01..FF60, U+FFE0..FFE6)
     \p{East_Asian_Width: H} \p{East_Asian_Width=Halfwidth} (123)
     \p{East_Asian_Width: Halfwidth} (Short: \p{Ea=H}) (123: U+20A9,
                          U+FF61..FFBE, U+FFC2..FFC7,
                          U+FFCA..FFCF, U+FFD2..FFD7, U+FFDA..FFDC
                          ...)
     \p{East_Asian_Width: N} \p{East_Asian_Width=Neutral} (792_699 plus
                          all above-Unicode code points)
     \p{East_Asian_Width: Na} \p{East_Asian_Width=Narrow} (111)
     \p{East_Asian_Width: Narrow} (Short: \p{Ea=Na}) (111: [\x20-\x7e
                          \xa2-\xa3\xa5-\xa6\xac\xaf],
                          U+27E6..27ED, U+2985..2986)
     \p{East_Asian_Width: Neutral} (Short: \p{Ea=N}) (792_699 plus all
                          above-Unicode code points: [\x00-\x1f
                          \x7f-\xa0\xa9\xab\xb5\xbb\xc0-\xc5\xc7-
                          \xcf\xd1-\xd6\xd9-\xdd\xe2-\xe5\xe7\xeb
                          \xee-\xef\xf1\xf4-\xf6\xfb\xfd\xff],
```

                    U+00FF..0100, U+0102..0110, U+0112,
                    U+0114..011A, U+011C..0125 ...)
      \p{East_Asian_Width: W} \p{East_Asian_Width=Wide} (182_336)
      \p{East_Asian_Width: Wide} (Short: \p{Ea=W}) (182_336:
                    U+1100..115F, U+231A..231B,
                    U+2329..232A, U+23E9..23EC, U+23F0,
                    U+23F3 ...)
      \p{EBase}   \p{Emoji_Modifier_Base} (=
                    \p{Emoji_Modifier_Base=Y}) (122)
      \p{EBase: *}   \p{Emoji_Modifier_Base: *}
      \p{EComp}   \p{Emoji_Component} (= \p{Emoji_Component=
                    Y}) (146)
      \p{EComp: *}   \p{Emoji_Component: *}
      \p{Egyp}   \p{Egyptian_Hieroglyphs} (=
                    \p{Script_Extensions=
                    Egyptian_Hieroglyphs}) (NOT \p{Block=
                    Egyptian_Hieroglyphs}) (1080)
   X \p{Egyptian_Hieroglyph_Format_Controls} \p{Block=
                    Egyptian_Hieroglyph_Format_Controls} (16)
    \p{Egyptian_Hieroglyphs} \p{Script_Extensions=
                    Egyptian_Hieroglyphs} (Short: \p{Egyp};
                    NOT \p{Block=Egyptian_Hieroglyphs})
                    (1080)
      \p{Elba}   \p{Elbasan} (= \p{Script_Extensions=
                    Elbasan}) (NOT \p{Block=Elbasan}) (40)
      \p{Elbasan}   \p{Script_Extensions=Elbasan} (Short:
                    \p{Elba}; NOT \p{Block=Elbasan}) (40)
      \p{Elym}   \p{Elymaic} (= \p{Script_Extensions=
                    Elymaic}) (NOT \p{Block=Elymaic}) (23)
      \p{Elymaic}   \p{Script_Extensions=Elymaic} (Short:
                    \p{Elym}; NOT \p{Block=Elymaic}) (23)
      \p{EMod}   \p{Emoji_Modifier} (= \p{Emoji_Modifier=
                    Y}) (5)
      \p{EMod: *}   \p{Emoji_Modifier: *}
      \p{Emoji}   \p{Emoji=Y} (1367)
      \p{Emoji: N*}   (Single: \P{Emoji}) (1_112_745 plus all
                    above-Unicode code points: [\x00-\x20!
                    \"\$\%&\'\'\(\)+,\-.\/:;<=>?\@A-Z\[\\\]
                    \^_`a-z\{\|\}~\x7f-\xa8\xaa-\xad\xaf-
                    \xff], U+0100..203B, U+203D..2048,

                    U+204A..2121, U+2123..2138, U+213A..2193
                    ...)
    \p{Emoji: Y*}   (Single: \p{Emoji}) (1367: [#*0-9\xa9
                    \xae], U+203C, U+2049, U+2122, U+2139,
                    U+2194..2199 ...)
    \p{Emoji_Component}   \p{Emoji_Component=Y} (Short: \p{EComp})
                    (146)
    \p{Emoji_Component: N*} (Short: \p{EComp=N}, \P{EComp}) (1_113_966
                    plus all above-Unicode code points:
                    [\x00-\x20!\"\$\%&\'\(\)+,\-.\/:;<=>?
                    \@A-Z\[\\\]\^_‘a-z\{\|\}~\x7f-\xff],
                    U+0100..200C, U+200E..20E2,
                    U+20E4..FE0E, U+FE10..1F1E5,
                    U+1F200..1F3FA ...)
    \p{Emoji_Component: Y*} (Short: \p{EComp=Y}, \p{EComp}) (146:
                    [#*0-9], U+200D, U+20E3, U+FE0F,
                    U+1F1E6..1F1FF, U+1F3FB..1F3FF ...)
    \p{Emoji_Modifier}   \p{Emoji_Modifier=Y} (Short: \p{EMod}) (5)
    \p{Emoji_Modifier: N*} (Short: \p{EMod=N}, \P{EMod}) (1_114_107
                    plus all above-Unicode code points:
                    U+0000..1F3FA, U+1F400..infinity)
    \p{Emoji_Modifier: Y*} (Short: \p{EMod=Y}, \p{EMod}) (5:
                    U+1F3FB..1F3FF)
    \p{Emoji_Modifier_Base} \p{Emoji_Modifier_Base=Y} (Short:
                    \p{EBase}) (122)
    \p{Emoji_Modifier_Base: N*} (Short: \p{EBase=N}, \P{EBase})
                    (1_113_990 plus all above-Unicode code
                    points: U+0000..261C, U+261E..26F8,
                    U+26FA..2709, U+270E..1F384,
                    U+1F386..1F3C1, U+1F3C5..1F3C6 ...)
    \p{Emoji_Modifier_Base: Y*} (Short: \p{EBase=Y}, \p{EBase}) (122:
                    U+261D, U+26F9, U+270A..270D, U+1F385,
                    U+1F3C2..1F3C4, U+1F3C7 ...)
    \p{Emoji_Presentation} \p{Emoji_Presentation=Y} (Short:
                    \p{EPres}) (1148)
    \p{Emoji_Presentation: N*} (Short: \p{EPres=N}, \P{EPres})
                    (1_112_964 plus all above-Unicode code
                    points: U+0000..2319, U+231C..23E8,
                    U+23ED..23EF, U+23F1..23F2,
                    U+23F4..25FC, U+25FF..2613 ...)

     \p{Emoji_Presentation: Y*} (Short: \p{EPres=Y}, \p{EPres}) (1148:
                    U+231A..231B, U+23E9..23EC, U+23F0,
                    U+23F3, U+25FD..25FE, U+2614..2615 ...)
X \p{Emoticons}   \p{Block=Emoticons} (80)
X \p{Enclosed_Alphanum}   \p{Enclosed_Alphanumerics} (= \p{Block=
                    Enclosed_Alphanumerics}) (160)
X \p{Enclosed_Alphanum_Sup} \p{Enclosed_Alphanumeric_Supplement} (=
                    \p{Block=
                    Enclosed_Alphanumeric_Supplement}) (256)
X \p{Enclosed_Alphanumeric_Supplement} \p{Block=
                    Enclosed_Alphanumeric_Supplement}
                    (Short: \p{InEnclosedAlphanumSup}) (256)
X \p{Enclosed_Alphanumerics} \p{Block=Enclosed_Alphanumerics}
                    (Short: \p{InEnclosedAlphanum}) (160)
X \p{Enclosed_CJK}   \p{Enclosed_CJK_Letters_And_Months} (=
                    \p{Block=
                    Enclosed_CJK_Letters_And_Months}) (256)
X \p{Enclosed_CJK_Letters_And_Months} \p{Block=
                    Enclosed_CJK_Letters_And_Months} (Short:
                    \p{InEnclosedCJK}) (256)
X \p{Enclosed_Ideographic_Sup} \p{Enclosed_Ideographic_Supplement}
                    (= \p{Block=
                    Enclosed_Ideographic_Supplement}) (256)
X \p{Enclosed_Ideographic_Supplement} \p{Block=
                    Enclosed_Ideographic_Supplement} (Short:
                    \p{InEnclosedIdeographicSup}) (256)
 \p{Enclosing_Mark}   \p{General_Category=Enclosing_Mark}
                    (Short: \p{Me}) (13)
 \p{EPres}   \p{Emoji_Presentation} (=
                    \p{Emoji_Presentation=Y}) (1148)
 \p{EPres: *}   \p{Emoji_Presentation: *}
 \p{Ethi}   \p{Ethiopic} (= \p{Script_Extensions=
                    Ethiopic}) (NOT \p{Block=Ethiopic}) (495)
 \p{Ethiopic}   \p{Script_Extensions=Ethiopic} (Short:
                    \p{Ethi}; NOT \p{Block=Ethiopic}) (495)
X \p{Ethiopic_Ext}   \p{Ethiopic_Extended} (= \p{Block=
                    Ethiopic_Extended}) (96)
X \p{Ethiopic_Ext_A}   \p{Ethiopic_Extended_A} (= \p{Block=
                    Ethiopic_Extended_A}) (48)
X \p{Ethiopic_Extended} \p{Block=Ethiopic_Extended} (Short:

```
                 \p{InEthiopicExt}) (96)
  X \p{Ethiopic_Extended_A} \p{Block=Ethiopic_Extended_A} (Short:
                 \p{InEthiopicExtA}) (48)
  X \p{Ethiopic_Sup}        \p{Ethiopic_Supplement} (= \p{Block=
                 Ethiopic_Supplement}) (32)
  X \p{Ethiopic_Supplement} \p{Block=Ethiopic_Supplement} (Short:
                 \p{InEthiopicSup}) (32)
   \p{Ext}             \p{Extender} (= \p{Extender=Y}) (48)
   \p{Ext: *}          \p{Extender: *}
   \p{Extended_Pictographic} \p{Extended_Pictographic=Y} (Short:
                 \p{ExtPict}) (3537)
   \p{Extended_Pictographic: N*} (Short: \p{ExtPict=N}, \P{ExtPict})
                 (1_110_575 plus all above-Unicode code
                 points: [\x00-\xa8\xaa-\xad\xaf-\xff],
                 U+0100..203B, U+203D..2048,
                 U+204A..2121, U+2123..2138, U+213A..2193
                 ...)
   \p{Extended_Pictographic: Y*} (Short: \p{ExtPict=Y}, \p{ExtPict})
                 (3537: [\xa9\xae], U+203C, U+2049,
                 U+2122, U+2139, U+2194..2199 ...)
   \p{Extender}        \p{Extender=Y} (Short: \p{Ext}) (48)
   \p{Extender: N*}    (Short: \p{Ext=N}, \P{Ext}) (1_114_064
                 plus all above-Unicode code points:
                 [\x00-\xb6\xb8-\xff], U+0100..02CF,
                 U+02D2..063F, U+0641..07F9,
                 U+07FB..0B54, U+0B56..0E45 ...)
   \p{Extender: Y*}    (Short: \p{Ext=Y}, \p{Ext}) (48: [\xb7],
                 U+02D0..02D1, U+0640, U+07FA, U+0B55,
                 U+0E46 ...)
   \p{ExtPict}         \p{Extended_Pictographic} (=
                 \p{Extended_Pictographic=Y}) (3537)
   \p{ExtPict: *}      \p{Extended_Pictographic: *}
   \p{Final_Punctuation}  \p{General_Category=Final_Punctuation}
                 (Short: \p{Pf}) (10)
   \p{Format}          \p{General_Category=Format} (Short:
                 \p{Cf}) (161)
   \p{Full_Composition_Exclusion} \p{Full_Composition_Exclusion=Y}
                 (Short: \p{CompEx}) (1120)
   \p{Full_Composition_Exclusion: N*} (Short: \p{CompEx=N},
                 \P{CompEx}) (1_112_992 plus all above-
```

          Unicode code points: U+0000..033F,
          U+0342, U+0345..0373, U+0375..037D,
          U+037F..0386, U+0388..0957 ...)

\p{Full_Composition_Exclusion: Y*} (Short: \p{CompEx=Y},
        \p{CompEx}) (1120: U+0340..0341,
        U+0343..0344, U+0374, U+037E, U+0387,
        U+0958..095F ...)

\p{Gc: *}   \p{General_Category: *}

\p{GCB: *}   \p{Grapheme_Cluster_Break: *}

\p{General_Category: C} \p{General_Category=Other} (970_414 plus
        all above-Unicode code points)

\p{General_Category: Cased_Letter} [\p{Ll}\p{Lu}\p{Lt}] (Short:
        \p{Gc=LC}, \p{LC}) (3977: [A-Za-z\xb5
        \xc0-\xd6\xd8-\xf6\xf8-\xff],
        U+0100..01BA, U+01BC..01BF,
        U+01C4..0293, U+0295..02AF, U+0370..0373
        ...)

\p{General_Category: Cc} \p{General_Category=Control} (65)

\p{General_Category: Cf} \p{General_Category=Format} (161)

\p{General_Category: Close_Punctuation} (Short: \p{Gc=Pe}, \p{Pe})
        (73: [\)\]\}], U+0F3B, U+0F3D, U+169C,
        U+2046, U+207E ...)

\p{General_Category: Cn} \p{General_Category=Unassigned} (830_672
        plus all above-Unicode code points)

\p{General_Category: Cntrl} \p{General_Category=Control} (65)

\p{General_Category: Co} \p{General_Category=Private_Use} (137_468)

\p{General_Category: Combining_Mark} \p{General_Category=Mark}
        (2295)

\p{General_Category: Connector_Punctuation} (Short: \p{Gc=Pc},
        \p{Pc}) (10: [_], U+203F..2040, U+2054,
        U+FE33..FE34, U+FE4D..FE4F, U+FF3F)

\p{General_Category: Control} (Short: \p{Gc=Cc}, \p{Cc}) (65:
        [\x00-\x1f\x7f-\x9f])

\p{General_Category: Cs} \p{General_Category=Surrogate} (2048)

\p{General_Category: Currency_Symbol} (Short: \p{Gc=Sc}, \p{Sc})
        (62: [\$\xa2-\xa5], U+058F, U+060B,
        U+07FE..07FF, U+09F2..09F3, U+09FB ...)

\p{General_Category: Dash_Punctuation} (Short: \p{Gc=Pd}, \p{Pd})
        (25: [\-], U+058A, U+05BE, U+1400,
        U+1806, U+2010..2015 ...)

\p{General_Category: Decimal_Number} (Short: \p{Gc=Nd}, \p{Nd})
                  (650: [0-9], U+0660..0669, U+06F0..06F9,
                  U+07C0..07C9, U+0966..096F, U+09E6..09EF
                  ...)
\p{General_Category: Digit} \p{General_Category=Decimal_Number}
                  (650)
\p{General_Category: Enclosing_Mark} (Short: \p{Gc=Me}, \p{Me})
                  (13: U+0488..0489, U+1ABE, U+20DD..20E0,
                  U+20E2..20E4, U+A670..A672)
\p{General_Category: Final_Punctuation} (Short: \p{Gc=Pf}, \p{Pf})
                  (10: [\xbb], U+2019, U+201D, U+203A,
                  U+2E03, U+2E05 ...)
\p{General_Category: Format} (Short: \p{Gc=Cf}, \p{Cf}) (161:
                  [\xad], U+0600..0605, U+061C, U+06DD,
                  U+070F, U+08E2 ...)
\p{General_Category: Initial_Punctuation} (Short: \p{Gc=Pi},
                  \p{Pi}) (12: [\xab], U+2018,
                  U+201B..201C, U+201F, U+2039, U+2E02 ...)
\p{General_Category: L} \p{General_Category=Letter} (131_241)
X \p{General_Category: L&} \p{General_Category=Cased_Letter} (3977)
X \p{General_Category: L_} \p{General_Category=Cased_Letter} Note
                  the trailing '_' matters in spite of
                  loose matching rules. (3977)
\p{General_Category: LC} \p{General_Category=Cased_Letter} (3977)
\p{General_Category: Letter} (Short: \p{Gc=L}, \p{L}) (131_241:
                  [A-Za-z\xaa\xb5\xba\xc0-\xd6\xd8-\xf6
                  \xf8-\xff], U+0100..02C1, U+02C6..02D1,
                  U+02E0..02E4, U+02EC, U+02EE ...)
\p{General_Category: Letter_Number} (Short: \p{Gc=Nl}, \p{Nl})
                  (236: U+16EE..16F0, U+2160..2182,
                  U+2185..2188, U+3007, U+3021..3029,
                  U+3038..303A ...)
\p{General_Category: Line_Separator} (Short: \p{Gc=Zl}, \p{Zl})
                  (1: U+2028)
\p{General_Category: Ll} \p{General_Category=Lowercase_Letter}
                  (/i= General_Category=Cased_Letter)
                  (2155)
\p{General_Category: Lm} \p{General_Category=Modifier_Letter} (260)
\p{General_Category: Lo} \p{General_Category=Other_Letter}
                  (127_004)

\p{General_Category: Lowercase_Letter} (Short: \p{Gc=Ll}, \p{Ll};
                /i= General_Category=Cased_Letter)
                (2155: [a-z\xb5\xdf-\xf6\xf8-\xff],
                U+0101, U+0103, U+0105, U+0107, U+0109
                ...)
\p{General_Category: Lt} \p{General_Category=Titlecase_Letter}
                (/i= General_Category=Cased_Letter) (31)
\p{General_Category: Lu} \p{General_Category=Uppercase_Letter}
                (/i= General_Category=Cased_Letter)
                (1791)
\p{General_Category: M} \p{General_Category=Mark} (2295)
\p{General_Category: Mark} (Short: \p{Gc=M}, \p{M}) (2295:
                U+0300..036F, U+0483..0489,
                U+0591..05BD, U+05BF, U+05C1..05C2,
                U+05C4..05C5 ...)
\p{General_Category: Math_Symbol} (Short: \p{Gc=Sm}, \p{Sm}) (948:
                [+<=>\|~\xac\xb1\xd7\xf7], U+03F6,
                U+0606..0608, U+2044, U+2052,
                U+207A..207C ...)
\p{General_Category: Mc} \p{General_Category=Spacing_Mark} (443)
\p{General_Category: Me} \p{General_Category=Enclosing_Mark} (13)
\p{General_Category: Mn} \p{General_Category=Nonspacing_Mark}
                (1839)
\p{General_Category: Modifier_Letter} (Short: \p{Gc=Lm}, \p{Lm})
                (260: U+02B0..02C1, U+02C6..02D1,
                U+02E0..02E4, U+02EC, U+02EE, U+0374 ...)
\p{General_Category: Modifier_Symbol} (Short: \p{Gc=Sk}, \p{Sk})
                (123: [\^`\xa8\xaf\xb4\xb8],
                U+02C2..02C5, U+02D2..02DF,
                U+02E5..02EB, U+02ED, U+02EF..02FF ...)
\p{General_Category: N} \p{General_Category=Number} (1781)
\p{General_Category: Nd} \p{General_Category=Decimal_Number} (650)
\p{General_Category: Nl} \p{General_Category=Letter_Number} (236)
\p{General_Category: No} \p{General_Category=Other_Number} (895)
\p{General_Category: Nonspacing_Mark} (Short: \p{Gc=Mn}, \p{Mn})
                (1839: U+0300..036F, U+0483..0487,
                U+0591..05BD, U+05BF, U+05C1..05C2,
                U+05C4..05C5 ...)
\p{General_Category: Number} (Short: \p{Gc=N}, \p{N}) (1781: [0-9
                \xb2-\xb3\xb9\xbc-\xbe], U+0660..0669,

                   U+06F0..06F9, U+07C0..07C9,
                   U+0966..096F, U+09E6..09EF ...)
\p{General_Category: Open_Punctuation} (Short: \p{Gc=Ps}, \p{Ps})
                   (75: [\(\[\{], U+0F3A, U+0F3C, U+169B,
                   U+201A, U+201E ...)
\p{General_Category: Other} (Short: \p{Gc=C}, \p{C}) (970_414 plus
                   all above-Unicode code points: [\x00-
                   \x1f\x7f-\x9f\xad], U+0378..0379,
                   U+0380..0383, U+038B, U+038D, U+03A2 ...)
\p{General_Category: Other_Letter} (Short: \p{Gc=Lo}, \p{Lo})
                   (127_004: [\xaa\xba], U+01BB,
                   U+01C0..01C3, U+0294, U+05D0..05EA,
                   U+05EF..05F2 ...)
\p{General_Category: Other_Number} (Short: \p{Gc=No}, \p{No})
                   (895: [\xb2-\xb3\xb9\xbc-\xbe],
                   U+09F4..09F9, U+0B72..0B77,
                   U+0BF0..0BF2, U+0C78..0C7E, U+0D58..0D5E
                   ...)
\p{General_Category: Other_Punctuation} (Short: \p{Gc=Po}, \p{Po})
                   (593: [!\"#\%&\'*,.\/:;?\@\\\xa1\xa7
                   \xb6-\xb7\xbf], U+037E, U+0387,
                   U+055A..055F, U+0589, U+05C0 ...)
\p{General_Category: Other_Symbol} (Short: \p{Gc=So}, \p{So})
                   (6431: [\xa6\xa9\xae\xb0], U+0482,
                   U+058D..058E, U+060E..060F, U+06DE,
                   U+06E9 ...)
\p{General_Category: P} \p{General_Category=Punctuation} (798)
\p{General_Category: Paragraph_Separator} (Short: \p{Gc=Zp},
                   \p{Zp}) (1: U+2029)
\p{General_Category: Pc} \p{General_Category=
                   Connector_Punctuation} (10)
\p{General_Category: Pd} \p{General_Category=Dash_Punctuation} (25)
\p{General_Category: Pe} \p{General_Category=Close_Punctuation}
                   (73)
\p{General_Category: Pf} \p{General_Category=Final_Punctuation}
                   (10)
\p{General_Category: Pi} \p{General_Category=Initial_Punctuation}
                   (12)
\p{General_Category: Po} \p{General_Category=Other_Punctuation}
                   (593)

\p{General_Category: Private_Use} (Short: \p{Gc=Co}, \p{Co})
        (137_468: U+E000..F8FF, U+F0000..FFFFD,
        U+100000..10FFFD)
\p{General_Category: Ps} \p{General_Category=Open_Punctuation} (75)
\p{General_Category: Punct} \p{General_Category=Punctuation} (798)
\p{General_Category: Punctuation} (Short: \p{Gc=P}, \p{P}) (798:
        [!\"#\%&\'\(\)*,\-.\/:;?\@\[\\\]_\{\}
        \xa1\xa7\xab\xb6-\xb7\xbb\xbf], U+037E,
        U+0387, U+055A..055F, U+0589..058A,
        U+05BE ...)
\p{General_Category: S} \p{General_Category=Symbol} (7564)
\p{General_Category: Sc} \p{General_Category=Currency_Symbol} (62)
\p{General_Category: Separator} (Short: \p{Gc=Z}, \p{Z}) (19:
        [\x20\xa0], U+1680, U+2000..200A,
        U+2028..2029, U+202F, U+205F ...)
\p{General_Category: Sk} \p{General_Category=Modifier_Symbol} (123)
\p{General_Category: Sm} \p{General_Category=Math_Symbol} (948)
\p{General_Category: So} \p{General_Category=Other_Symbol} (6431)
\p{General_Category: Space_Separator} (Short: \p{Gc=Zs}, \p{Zs})
        (17: [\x20\xa0], U+1680, U+2000..200A,
        U+202F, U+205F, U+3000)
\p{General_Category: Spacing_Mark} (Short: \p{Gc=Mc}, \p{Mc})
        (443: U+0903, U+093B, U+093E..0940,
        U+0949..094C, U+094E..094F, U+0982..0983
        ...)
\p{General_Category: Surrogate} (Short: \p{Gc=Cs}, \p{Cs}) (2048:
        U+D800..DFFF)
\p{General_Category: Symbol} (Short: \p{Gc=S}, \p{S}) (7564:
        [\$+<=>\^\`\|~\xa2-\xa6\xa8-\xa9\xac\xae-
        \xb1\xb4\xb8\xd7\xf7], U+02C2..02C5,
        U+02D2..02DF, U+02E5..02EB, U+02ED,
        U+02EF..02FF ...)
\p{General_Category: Titlecase_Letter} (Short: \p{Gc=Lt}, \p{Lt};
        /i= General_Category=Cased_Letter) (31:
        U+01C5, U+01C8, U+01CB, U+01F2,
        U+1F88..1F8F, U+1F98..1F9F ...)
\p{General_Category: Unassigned} (Short: \p{Gc=Cn}, \p{Cn})
        (830_672 plus all above-Unicode code
        points: U+0378..0379, U+0380..0383,
        U+038B, U+038D, U+03A2, U+0530 ...)

\p{General_Category: Uppercase_Letter} (Short: \p{Gc=Lu}, \p{Lu};
              /i= General_Category=Cased_Letter)
              (1791: [A-Z\xc0-\xd6\xd8-\xde], U+0100,
              U+0102, U+0104, U+0106, U+0108 ...)
\p{General_Category: Z} \p{General_Category=Separator} (19)
\p{General_Category: Zl} \p{General_Category=Line_Separator} (1)
\p{General_Category: Zp} \p{General_Category=Paragraph_Separator}
              (1)
\p{General_Category: Zs} \p{General_Category=Space_Separator} (17)
X \p{General_Punctuation} \p{Block=General_Punctuation} (Short:
              \p{InPunctuation}) (112)
X \p{Geometric_Shapes}   \p{Block=Geometric_Shapes} (96)
X \p{Geometric_Shapes_Ext} \p{Geometric_Shapes_Extended} (=
              \p{Block=Geometric_Shapes_Extended})
              (128)
X \p{Geometric_Shapes_Extended} \p{Block=Geometric_Shapes_Extended}
              (Short: \p{InGeometricShapesExt}) (128)
 \p{Geor}   \p{Georgian} (= \p{Script_Extensions=
              Georgian}) (NOT \p{Block=Georgian}) (174)
 \p{Georgian}   \p{Script_Extensions=Georgian} (Short:
              \p{Geor}; NOT \p{Block=Georgian}) (174)
X \p{Georgian_Ext}   \p{Georgian_Extended} (= \p{Block=
              Georgian_Extended}) (48)
X \p{Georgian_Extended}   \p{Block=Georgian_Extended} (Short:
              \p{InGeorgianExt}) (48)
X \p{Georgian_Sup}   \p{Georgian_Supplement} (= \p{Block=
              Georgian_Supplement}) (48)
X \p{Georgian_Supplement} \p{Block=Georgian_Supplement} (Short:
              \p{InGeorgianSup}) (48)
 \p{Glag}   \p{Glagolitic} (= \p{Script_Extensions=
              Glagolitic}) (NOT \p{Block=Glagolitic})
              (136)
 \p{Glagolitic}   \p{Script_Extensions=Glagolitic} (Short:
              \p{Glag}; NOT \p{Block=Glagolitic}) (136)
X \p{Glagolitic_Sup}   \p{Glagolitic_Supplement} (= \p{Block=
              Glagolitic_Supplement}) (48)
X \p{Glagolitic_Supplement} \p{Block=Glagolitic_Supplement} (Short:
              \p{InGlagoliticSup}) (48)
 \p{Gong}   \p{Gunjala_Gondi} (= \p{Script_Extensions=
              Gunjala_Gondi}) (NOT \p{Block=

```
                  Gunjala_Gondi}) (65)
     \p{Gonm}          \p{Masaram_Gondi} (= \p{Script_Extensions=
                  Masaram_Gondi}) (NOT \p{Block=
                  Masaram_Gondi}) (77)
     \p{Goth}          \p{Gothic} (= \p{Script_Extensions=
                  Gothic}) (NOT \p{Block=Gothic}) (27)
     \p{Gothic}         \p{Script_Extensions=Gothic} (Short:
                  \p{Goth}; NOT \p{Block=Gothic}) (27)
     \p{Gr_Base}         \p{Grapheme_Base} (= \p{Grapheme_Base=Y})
                  (141_814)
     \p{Gr_Base: *}       \p{Grapheme_Base: *}
     \p{Gr_Ext}          \p{Grapheme_Extend} (= \p{Grapheme_Extend=
                  Y}) (1979)
     \p{Gr_Ext: *}        \p{Grapheme_Extend: *}
     \p{Gran}           \p{Grantha} (= \p{Script_Extensions=
                  Grantha}) (NOT \p{Block=Grantha}) (116)
     \p{Grantha}         \p{Script_Extensions=Grantha} (Short:
                  \p{Gran}; NOT \p{Block=Grantha}) (116)
     \p{Graph}          \p{XPosixGraph} (281_308)
     \p{Grapheme_Base}     \p{Grapheme_Base=Y} (Short: \p{GrBase})
                  (141_814)
     \p{Grapheme_Base: N*}   (Short: \p{GrBase=N}, \P{GrBase}) (972_298
                  plus all above-Unicode code points:
                  [\x00-\x1f\x7f-\x9f\xad], U+0300..036F,
                  U+0378..0379, U+0380..0383, U+038B,
                  U+038D ...)
     \p{Grapheme_Base: Y*}   (Short: \p{GrBase=Y}, \p{GrBase})
                  (141_814: [\x20-\x7e\xa0-\xac\xae-\xff],
                  U+0100..02FF, U+0370..0377,
                  U+037A..037F, U+0384..038A, U+038C ...)
     \p{Grapheme_Cluster_Break: CN} \p{Grapheme_Cluster_Break=Control}
                  (3886)
     \p{Grapheme_Cluster_Break: Control} (Short: \p{GCB=CN}) (3886: [^
                  \n\r\x20-\x7e\xa0-\xac\xae-\xff],
                  U+061C, U+180E, U+200B, U+200E..200F,
                  U+2028..202E ...)
     \p{Grapheme_Cluster_Break: CR} (Short: \p{GCB=CR}) (1: [\r])
     \p{Grapheme_Cluster_Break: E_Base} (Short: \p{GCB=EB}) (0)
     \p{Grapheme_Cluster_Break: E_Base_GAZ} (Short: \p{GCB=EBG}) (0)
     \p{Grapheme_Cluster_Break: E_Modifier} (Short: \p{GCB=EM}) (0)
```

\p{Grapheme_Cluster_Break: EB} \p{Grapheme_Cluster_Break=E_Base}
                (0)
\p{Grapheme_Cluster_Break: EBG} \p{Grapheme_Cluster_Break=
                E_Base_GAZ} (0)
\p{Grapheme_Cluster_Break: EM} \p{Grapheme_Cluster_Break=
                E_Modifier} (0)
\p{Grapheme_Cluster_Break: EX} \p{Grapheme_Cluster_Break=Extend}
                (1984)
\p{Grapheme_Cluster_Break: Extend} (Short: \p{GCB=EX}) (1984:
                U+0300..036F, U+0483..0489,
                U+0591..05BD, U+05BF, U+05C1..05C2,
                U+05C4..05C5 ...)
\p{Grapheme_Cluster_Break: GAZ} \p{Grapheme_Cluster_Break=
                Glue_After_Zwj} (0)
\p{Grapheme_Cluster_Break: Glue_After_Zwj} (Short: \p{GCB=GAZ}) (0)
\p{Grapheme_Cluster_Break: L} (Short: \p{GCB=L}) (125:
                U+1100..115F, U+A960..A97C)
\p{Grapheme_Cluster_Break: LF} (Short: \p{GCB=LF}) (1: [\n])
\p{Grapheme_Cluster_Break: LV} (Short: \p{GCB=LV}) (399: U+AC00,
                U+AC1C, U+AC38, U+AC54, U+AC70, U+AC8C
                ...)
\p{Grapheme_Cluster_Break: LVT} (Short: \p{GCB=LVT}) (10_773:
                U+AC01..AC1B, U+AC1D..AC37,
                U+AC39..AC53, U+AC55..AC6F,
                U+AC71..AC8B, U+AC8D..ACA7 ...)
\p{Grapheme_Cluster_Break: Other} (Short: \p{GCB=XX}) (1_096_272
                plus all above-Unicode code points:
                [\x20-\x7e\xa0-\xac\xae-\xff],
                U+0100..02FF, U+0370..0482,
                U+048A..0590, U+05BE, U+05C0 ...)
\p{Grapheme_Cluster_Break: PP} \p{Grapheme_Cluster_Break=Prepend}
                (24)
\p{Grapheme_Cluster_Break: Prepend} (Short: \p{GCB=PP}) (24:
                U+0600..0605, U+06DD, U+070F, U+08E2,
                U+0D4E, U+110BD ...)
\p{Grapheme_Cluster_Break: Regional_Indicator} (Short: \p{GCB=RI})
                (26: U+1F1E6..1F1FF)
\p{Grapheme_Cluster_Break: RI} \p{Grapheme_Cluster_Break=
                Regional_Indicator} (26)
\p{Grapheme_Cluster_Break: SM} \p{Grapheme_Cluster_Break=

SpacingMark} (388)
\p{Grapheme_Cluster_Break: SpacingMark} (Short: \p{GCB=SM}) (388:
        U+0903, U+093B, U+093E..0940,
        U+0949..094C, U+094E..094F, U+0982..0983
        ...)
\p{Grapheme_Cluster_Break: T} (Short: \p{GCB=T}) (137:
        U+11A8..11FF, U+D7CB..D7FB)
\p{Grapheme_Cluster_Break: V} (Short: \p{GCB=V}) (95:
        U+1160..11A7, U+D7B0..D7C6)
\p{Grapheme_Cluster_Break: XX} \p{Grapheme_Cluster_Break=Other}
        (1_096_272 plus all above-Unicode code
        points)
\p{Grapheme_Cluster_Break: ZWJ} (Short: \p{GCB=ZWJ}) (1: U+200D)
\p{Grapheme_Extend}   \p{Grapheme_Extend=Y} (Short: \p{GrExt})
        (1979)
\p{Grapheme_Extend: N*} (Short: \p{GrExt=N}, \P{GrExt}) (1_112_133
        plus all above-Unicode code points:
        U+0000..02FF, U+0370..0482,
        U+048A..0590, U+05BE, U+05C0, U+05C3 ...)
\p{Grapheme_Extend: Y*} (Short: \p{GrExt=Y}, \p{GrExt}) (1979:
        U+0300..036F, U+0483..0489,
        U+0591..05BD, U+05BF, U+05C1..05C2,
        U+05C4..05C5 ...)
\p{Greek}   \p{Script_Extensions=Greek} (Short:
        \p{Grek}; NOT \p{Greek_And_Coptic}) (522)
X \p{Greek_And_Coptic}   \p{Block=Greek_And_Coptic} (Short:
        \p{InGreek}) (144)
X \p{Greek_Ext}   \p{Greek_Extended} (= \p{Block=
        Greek_Extended}) (256)
X \p{Greek_Extended}   \p{Block=Greek_Extended} (Short:
        \p{InGreekExt}) (256)
\p{Grek}   \p{Greek} (= \p{Script_Extensions=Greek})
        (NOT \p{Greek_And_Coptic}) (522)
\p{Gujarati}   \p{Script_Extensions=Gujarati} (Short:
        \p{Gujr}; NOT \p{Block=Gujarati}) (105)
\p{Gujr}   \p{Gujarati} (= \p{Script_Extensions=
        Gujarati}) (NOT \p{Block=Gujarati}) (105)
\p{Gunjala_Gondi}   \p{Script_Extensions=Gunjala_Gondi}
        (Short: \p{Gong}; NOT \p{Block=
        Gunjala_Gondi}) (65)

   \p{Gurmukhi}   \p{Script_Extensions=Gurmukhi} (Short:
             \p{Guru}; NOT \p{Block=Gurmukhi}) (94)

  \p{Guru}   \p{Gurmukhi} (= \p{Script_Extensions=
             Gurmukhi}) (NOT \p{Block=Gurmukhi}) (94)

X \p{Half_And_Full_Forms} \p{Halfwidth_And_Fullwidth_Forms} (=
             \p{Block=Halfwidth_And_Fullwidth_Forms})
             (240)

X \p{Half_Marks}   \p{Combining_Half_Marks} (= \p{Block=
             Combining_Half_Marks}) (16)

X \p{Halfwidth_And_Fullwidth_Forms} \p{Block=
             Halfwidth_And_Fullwidth_Forms} (Short:
             \p{InHalfAndFullForms}) (240)

 \p{Han}   \p{Script_Extensions=Han} (94_492)

 \p{Hang}   \p{Hangul} (= \p{Script_Extensions=
             Hangul}) (NOT \p{Hangul_Syllables})
             (11_775)

 \p{Hangul}   \p{Script_Extensions=Hangul} (Short:
             \p{Hang}; NOT \p{Hangul_Syllables})
             (11_775)

X \p{Hangul_Compatibility_Jamo} \p{Block=Hangul_Compatibility_Jamo}
             (Short: \p{InCompatJamo}) (96)

X \p{Hangul_Jamo}   \p{Block=Hangul_Jamo} (Short: \p{InJamo})
             (256)

X \p{Hangul_Jamo_Extended_A} \p{Block=Hangul_Jamo_Extended_A}
             (Short: \p{InJamoExtA}) (32)

X \p{Hangul_Jamo_Extended_B} \p{Block=Hangul_Jamo_Extended_B}
             (Short: \p{InJamoExtB}) (80)

 \p{Hangul_Syllable_Type: L} \p{Hangul_Syllable_Type=Leading_Jamo}
             (125)

 \p{Hangul_Syllable_Type: Leading_Jamo} (Short: \p{Hst=L}) (125:
             U+1100..115F, U+A960..A97C)

 \p{Hangul_Syllable_Type: LV} \p{Hangul_Syllable_Type=LV_Syllable}
             (399)

 \p{Hangul_Syllable_Type: LV_Syllable} (Short: \p{Hst=LV}) (399:
             U+AC00, U+AC1C, U+AC38, U+AC54, U+AC70,
             U+AC8C ...)

 \p{Hangul_Syllable_Type: LVT} \p{Hangul_Syllable_Type=
             LVT_Syllable} (10_773)

 \p{Hangul_Syllable_Type: LVT_Syllable} (Short: \p{Hst=LVT})
             (10_773: U+AC01..AC1B, U+AC1D..AC37,

                      U+AC39..AC53, U+AC55..AC6F,
                      U+AC71..AC8B, U+AC8D..ACA7 ...)
   \p{Hangul_Syllable_Type: NA} \p{Hangul_Syllable_Type=
                      Not_Applicable} (1_102_583 plus all
                      above-Unicode code points)
   \p{Hangul_Syllable_Type: Not_Applicable} (Short: \p{Hst=NA})
                      (1_102_583 plus all above-Unicode code
                      points: U+0000..10FF, U+1200..A95F,
                      U+A97D..ABFF, U+D7A4..D7AF,
                      U+D7C7..D7CA, U+D7FC..infinity)
   \p{Hangul_Syllable_Type: T} \p{Hangul_Syllable_Type=Trailing_Jamo}
                      (137)
   \p{Hangul_Syllable_Type: Trailing_Jamo} (Short: \p{Hst=T}) (137:
                      U+11A8..11FF, U+D7CB..D7FB)
   \p{Hangul_Syllable_Type: V} \p{Hangul_Syllable_Type=Vowel_Jamo}
                      (95)
   \p{Hangul_Syllable_Type: Vowel_Jamo} (Short: \p{Hst=V}) (95:
                      U+1160..11A7, U+D7B0..D7C6)
 X \p{Hangul_Syllables}   \p{Block=Hangul_Syllables} (Short:
                      \p{InHangul}) (11_184)
   \p{Hani}   \p{Han} (= \p{Script_Extensions=Han})
                      (94_492)
   \p{Hanifi_Rohingya}   \p{Script_Extensions=Hanifi_Rohingya}
                      (Short: \p{Rohg}; NOT \p{Block=
                      Hanifi_Rohingya}) (55)
   \p{Hano}   \p{Hanunoo} (= \p{Script_Extensions=
                      Hanunoo}) (NOT \p{Block=Hanunoo}) (23)
   \p{Hanunoo}   \p{Script_Extensions=Hanunoo} (Short:
                      \p{Hano}; NOT \p{Block=Hanunoo}) (23)
   \p{Hatr}   \p{Hatran} (= \p{Script_Extensions=
                      Hatran}) (NOT \p{Block=Hatran}) (26)
   \p{Hatran}   \p{Script_Extensions=Hatran} (Short:
                      \p{Hatr}; NOT \p{Block=Hatran}) (26)
   \p{Hebr}   \p{Hebrew} (= \p{Script_Extensions=
                      Hebrew}) (NOT \p{Block=Hebrew}) (134)
   \p{Hebrew}   \p{Script_Extensions=Hebrew} (Short:
                      \p{Hebr}; NOT \p{Block=Hebrew}) (134)
   \p{Hex}   \p{XPosixXDigit} (= \p{Hex_Digit=Y}) (44)
   \p{Hex: *}   \p{Hex_Digit: *}
   \p{Hex_Digit}   \p{XPosixXDigit} (= \p{Hex_Digit=Y}) (44)

\p{Hex_Digit: N*}   (Short: \p{Hex=N}, \P{Hex}) (1_114_068
                 plus all above-Unicode code points:
                 [\x00-\x20!\"#\$\%&\'\(\)*+,\-.\/:;<=>?
                 \@G-Z\[\\\]\^_‘g-z\{\|\}~\x7f-\xff],
                 U+0100..FF0F, U+FF1A..FF20,
                 U+FF27..FF40, U+FF47..infinity)
\p{Hex_Digit: Y*}   (Short: \p{Hex=Y}, \p{Hex}) (44: [0-9A-Fa-
                 f], U+FF10..FF19, U+FF21..FF26,
                 U+FF41..FF46)
X \p{High_Private_Use_Surrogates} \p{Block=
                 High_Private_Use_Surrogates} (Short:
                 \p{InHighPUSurrogates}) (128)
X \p{High_PU_Surrogates} \p{High_Private_Use_Surrogates} (=
                 \p{Block=High_Private_Use_Surrogates})
                 (128)
X \p{High_Surrogates}   \p{Block=High_Surrogates} (896)
 \p{Hira}   \p{Hiragana} (= \p{Script_Extensions=
                 Hiragana}) (NOT \p{Block=Hiragana}) (431)
 \p{Hiragana}   \p{Script_Extensions=Hiragana} (Short:
                 \p{Hira}; NOT \p{Block=Hiragana}) (431)
 \p{Hluw}   \p{Anatolian_Hieroglyphs} (=
                 \p{Script_Extensions=
                 Anatolian_Hieroglyphs}) (NOT \p{Block=
                 Anatolian_Hieroglyphs}) (583)
 \p{Hmng}   \p{Pahawh_Hmong} (= \p{Script_Extensions=
                 Pahawh_Hmong}) (NOT \p{Block=
                 Pahawh_Hmong}) (127)
 \p{Hmnp}   \p{Nyiakeng_Puachue_Hmong} (=
                 \p{Script_Extensions=
                 Nyiakeng_Puachue_Hmong}) (NOT \p{Block=
                 Nyiakeng_Puachue_Hmong}) (71)
 \p{HorizSpace}   \p{XPosixBlank} (18)
 \p{Hst: *}   \p{Hangul_Syllable_Type: *}
 \p{Hung}   \p{Old_Hungarian} (= \p{Script_Extensions=
                 Old_Hungarian}) (NOT \p{Block=
                 Old_Hungarian}) (108)
D \p{Hyphen}   \p{Hyphen=Y} (11)
D \p{Hyphen: N*}   Supplanted by Line_Break property values;
                 see www.unicode.org/reports/tr14
                 (Single: \P{Hyphen}) (1_114_101 plus all

                above-Unicode code points: [\x00-\x20!
                \"#\$\%&\'\(\)*+,.\/0-9:;<=>?\@A-Z
                \[\\\]\^_'a-z\{\|\}~\x7f-\xac\xae-\xff],
                U+0100..0589, U+058B..1805,
                U+1807..200F, U+2012..2E16, U+2E18..30FA
                ...)

D \p{Hyphen: Y*}   Supplanted by Line_Break property values;
                see www.unicode.org/reports/tr14
                (Single: \p{Hyphen}) (11: [\-\xad],
                U+058A, U+1806, U+2010..2011, U+2E17,
                U+30FB ...)

\p{ID_Continue}   \p{ID_Continue=Y} (Short: \p{IDC}; NOT
                \p{Ideographic_Description_Characters})
                (134_434)

\p{ID_Continue: N*}   (Short: \p{IDC=N}, \P{IDC}) (979_678 plus
                all above-Unicode code points: [\x00-
                \x20!\"#\$\%&\'\(\)*+,\-.\/:;<=>?\@
                \[\\\]\^'\{\|\}~\x7f-\xa9\xab-\xb4\xb6
                \xb8-\xb9\xbb-\xbf\xd7\xf7],
                U+02C2..02C5, U+02D2..02DF,
                U+02E5..02EB, U+02ED, U+02EF..02FF ...)

\p{ID_Continue: Y*}   (Short: \p{IDC=Y}, \p{IDC}) (134_434:
                [0-9A-Z_a-z\xaa\xb5\xb7\xba\xc0-\xd6
                \xd8-\xf6\xf8-\xff], U+0100..02C1,
                U+02C6..02D1, U+02E0..02E4, U+02EC,
                U+02EE ...)

\p{ID_Start}   \p{ID_Start=Y} (Short: \p{IDS}) (131_482)

\p{ID_Start: N*}   (Short: \p{IDS=N}, \P{IDS}) (982_630 plus
                all above-Unicode code points: [\x00-
                \x20!\"#\$\%&\'\(\)*+,\-.\/0-9:;<=>?\@
                \[\\\]\^_'\{\|\}~\x7f-\xa9\xab-\xb4\xb6-
                \xb9\xbb-\xbf\xd7\xf7], U+02C2..02C5,
                U+02D2..02DF, U+02E5..02EB, U+02ED,
                U+02EF..036F ...)

\p{ID_Start: Y*}   (Short: \p{IDS=Y}, \p{IDS}) (131_482: [A-
                Za-z\xaa\xb5\xba\xc0-\xd6\xd8-\xf6\xf8-
                \xff], U+0100..02C1, U+02C6..02D1,
                U+02E0..02E4, U+02EC, U+02EE ...)

\p{IDC}   \p{ID_Continue} (= \p{ID_Continue=Y}) (NOT
                \p{Ideographic_Description_Characters})

          (134_434)
\p{IDC: *}   \p{ID_Continue: *}
\p{Identifier_Status: Allowed} (107_835: [\'\-.0-9:A-Z_a-z\xb7
                \xc0-\xd6\xd8-\xf6\xf8-\xff],
                U+0100..0131, U+0134..013E,
                U+0141..0148, U+014A..017E, U+018F ...)
\p{Identifier_Status: Restricted} (1_006_277 plus all above-
                Unicode code points: [\x00-\x20!\"#\$
                \%&\(\)*+,\/;<=>?\@\[\\\]\^\`\{\|\}~\x7f-
                \xb6\xb8-\xbf\xd7\xf7], U+0132..0133,
                U+013F..0140, U+0149, U+017F..018E,
                U+0190..019F ...)
\p{Identifier_Type: Default_Ignorable} (395: [\xad], U+034F,
                U+061C, U+115F..1160, U+17B4..17B5,
                U+180B..180E ...)
\p{Identifier_Type: Deprecated} (15: U+0149, U+0673, U+0F77,
                U+0F79, U+17A3..17A4, U+206A..206F ...)
\p{Identifier_Type: Exclusion} (16_745: U+03E2..03EF,
                U+0800..082D, U+0830..083E,
                U+1680..169C, U+16A0..16EA, U+16EE..16F8
                ...)
\p{Identifier_Type: Inclusion} (19: [\'\-.:\xb7], U+0375, U+058A,
                U+05F3..05F4, U+06FD..06FE, U+0F0B ...)
\p{Identifier_Type: Limited_Use} (5248: U+0700..070D,
                U+070F..074A, U+074D..074F,
                U+07C0..07FA, U+07FD..07FF, U+0840..085B
                ...)
\p{Identifier_Type: Not_Character} (970_247 plus all above-Unicode
                code points: [^\t\n\cK\f\r\x20-\x7e\x85
                \xa0-\xff], U+0378..0379, U+0380..0383,
                U+038B, U+038D, U+03A2 ...)
\p{Identifier_Type: Not_NFKC} (4800: [\xa0\xa8\xaa\xaf\xb2-\xb5
                \xb8-\xba\xbc-\xbe], U+0132..0133,
                U+013F..0140, U+017F, U+01C4..01CC,
                U+01F1..01F3 ...)
\p{Identifier_Type: Not_XID} (7998: [\t\n\cK\f\r\x20!\"#\$\%&
                \(\)*+,\/;<=>?\@\[\\\]\^\`\{\|\}~\x85
                \xa1-\xa7\xa9\xab-\xac\xae\xb0-\xb1\xb6
                \xbb\xbf\xd7\xf7], U+02C2..02C5,
                U+02D2..02D7, U+02DE..02DF,

U+02E5..02EB, U+02ED ...)
\p{Identifier_Type: Obsolete} (1611: U+018D, U+01AA..01AB,
                  U+01B9..01BB, U+01BE..01BF,
                  U+01F6..01F7, U+021C..021D ...)
\p{Identifier_Type: Recommended} (107_816: [0-9A-Z_a-z\xc0-\xd6
                  \xd8-\xf6\xf8-\xff], U+0100..0131,
                  U+0134..013E, U+0141..0148,
                  U+014A..017E, U+018F ...)
\p{Identifier_Type: Technical} (1463: U+0180, U+018D,
                  U+01AA..01AB, U+01BA..01BB, U+01BE,
                  U+01C0..01C3 ...)
\p{Identifier_Type: Uncommon_Use} (348: U+0181..018C, U+018E,
                  U+0190..019F, U+01A2..01A9,
                  U+01AC..01AE, U+01B1..01B8 ...)
\p{Ideo}   \p{Ideographic} (= \p{Ideographic=Y})
                  (101_652)
\p{Ideo: *}   \p{Ideographic: *}
\p{Ideographic}   \p{Ideographic=Y} (Short: \p{Ideo})
                  (101_652)
\p{Ideographic: N*}   (Short: \p{Ideo=N}, \P{Ideo}) (1_012_460
                  plus all above-Unicode code points:
                  U+0000..3005, U+3008..3020,
                  U+302A..3037, U+303B..33FF,
                  U+4DC0..4DFF, U+9FFD..F8FF ...)
\p{Ideographic: Y*}   (Short: \p{Ideo=Y}, \p{Ideo}) (101_652:
                  U+3006..3007, U+3021..3029,
                  U+3038..303A, U+3400..4DBF,
                  U+4E00..9FFC, U+F900..FA6D ...)
X \p{Ideographic_Description_Characters} \p{Block=
                  Ideographic_Description_Characters}
                  (Short: \p{InIDC}) (16)
X \p{Ideographic_Symbols} \p{Ideographic_Symbols_And_Punctuation} (=
                  \p{Block=
                  Ideographic_Symbols_And_Punctuation})
                  (32)
X \p{Ideographic_Symbols_And_Punctuation} \p{Block=
                  Ideographic_Symbols_And_Punctuation}
                  (Short: \p{InIdeographicSymbols}) (32)
\p{IDS}   \p{ID_Start} (= \p{ID_Start=Y}) (131_482)
\p{IDS: *}   \p{ID_Start: *}

\p{IDS_Binary_Operator} \p{IDS_Binary_Operator=Y} (Short:
                \p{IDSB}) (10)

\p{IDS_Binary_Operator: N*} (Short: \p{IDSB=N}, \P{IDSB})
                (1_114_102 plus all above-Unicode code
                points: U+0000..2FEF, U+2FF2..2FF3,
                U+2FFC..infinity)

\p{IDS_Binary_Operator: Y*} (Short: \p{IDSB=Y}, \p{IDSB}) (10:
                U+2FF0..2FF1, U+2FF4..2FFB)

\p{IDS_Trinary_Operator} \p{IDS_Trinary_Operator=Y} (Short:
                \p{IDST}) (2)

\p{IDS_Trinary_Operator: N*} (Short: \p{IDST=N}, \P{IDST})
                (1_114_110 plus all above-Unicode code
                points: U+0000..2FF1, U+2FF4..infinity)

\p{IDS_Trinary_Operator: Y*} (Short: \p{IDST=Y}, \p{IDST}) (2:
                U+2FF2..2FF3)

\p{IDSB}   \p{IDS_Binary_Operator} (=
                \p{IDS_Binary_Operator=Y}) (10)

\p{IDSB: *}   \p{IDS_Binary_Operator: *}

\p{IDST}   \p{IDS_Trinary_Operator} (=
                \p{IDS_Trinary_Operator=Y}) (2)

\p{IDST: *}   \p{IDS_Trinary_Operator: *}

\p{Imperial_Aramaic} \p{Script_Extensions=Imperial_Aramaic}
                (Short: \p{Armi}; NOT \p{Block=
                Imperial_Aramaic}) (31)

\p{In: *}   \p{Present_In: *} (Perl extension)

X \p{In_*}   \p{Block: *}

X \p{Indic_Number_Forms} \p{Common_Indic_Number_Forms} (= \p{Block=
                Common_Indic_Number_Forms}) (16)

\p{Indic_Positional_Category: Bottom} (Short: \p{InPC=Bottom})
                (351: U+093C, U+0941..0944, U+094D,
                U+0952, U+0956..0957, U+0962..0963 ...)

\p{Indic_Positional_Category: Bottom_And_Left} (Short: \p{InPC=
                BottomAndLeft}) (1: U+A9BF)

\p{Indic_Positional_Category: Bottom_And_Right} (Short: \p{InPC=
                BottomAndRight}) (4: U+1B3B, U+A9BE,
                U+A9C0, U+11942)

\p{Indic_Positional_Category: Left} (Short: \p{InPC=Left}) (64:
                U+093F, U+094E, U+09BF, U+09C7..09C8,
                U+0A3F, U+0ABF ...)

\p{Indic_Positional_Category: Left_And_Right} (Short: \p{InPC=

LeftAndRight}) (22: U+09CB..09CC,
U+0B4B, U+0BCA..0BCC, U+0D4A..0D4C,
U+0DDC, U+0DDE ...)
\p{Indic_Positional_Category: NA} (Short: \p{InPC=NA}) (1_112_902
plus all above-Unicode code points:
U+0000..08FF, U+0904..0939, U+093D,
U+0950, U+0958..0961, U+0964..0980 ...)
\p{Indic_Positional_Category: Overstruck} (Short: \p{InPC=
Overstruck}) (10: U+1CD4, U+1CE2..1CE8,
U+10A01, U+10A06)
\p{Indic_Positional_Category: Right} (Short: \p{InPC=Right}) (288:
U+0903, U+093B, U+093E, U+0940,
U+0949..094C, U+094F ...)
\p{Indic_Positional_Category: Top} (Short: \p{InPC=Top}) (415:
U+0900..0902, U+093A, U+0945..0948,
U+0951, U+0953..0955, U+0981 ...)
\p{Indic_Positional_Category: Top_And_Bottom} (Short: \p{InPC=
TopAndBottom}) (10: U+0C48, U+0F73,
U+0F76..0F79, U+0F81, U+1B3C,
U+1112E..1112F)
\p{Indic_Positional_Category: Top_And_Bottom_And_Left} (Short:
\p{InPC=TopAndBottomAndLeft}) (2:
U+103C, U+1171E)
\p{Indic_Positional_Category: Top_And_Bottom_And_Right} (Short:
\p{InPC=TopAndBottomAndRight}) (1:
U+1B3D)
\p{Indic_Positional_Category: Top_And_Left} (Short: \p{InPC=
TopAndLeft}) (6: U+0B48, U+0DDA, U+17BE,
U+1C29, U+114BB, U+115B9)
\p{Indic_Positional_Category: Top_And_Left_And_Right} (Short:
\p{InPC=TopAndLeftAndRight}) (4: U+0B4C,
U+0DDD, U+17BF, U+115BB)
\p{Indic_Positional_Category: Top_And_Right} (Short: \p{InPC=
TopAndRight}) (13: U+0AC9, U+0B57,
U+0CC0, U+0CC7..0CC8, U+0CCA..0CCB,
U+1925..1926 ...)
\p{Indic_Positional_Category: Visual_Order_Left} (Short: \p{InPC=
VisualOrderLeft}) (19: U+0E40..0E44,
U+0EC0..0EC4, U+19B5..19B7, U+19BA,
U+AAB5..AAB6, U+AAB9 ...)

X \p{Indic_Siyaq_Numbers} \p{Block=Indic_Siyaq_Numbers} (80)
  \p{Indic_Syllabic_Category: Avagraha} (Short: \p{InSC=Avagraha})
                (17: U+093D, U+09BD, U+0ABD, U+0B3D,
                U+0C3D, U+0CBD ...)
  \p{Indic_Syllabic_Category: Bindu} (Short: \p{InSC=Bindu}) (91:
                U+0900..0902, U+0981..0982, U+09FC,
                U+0A01..0A02, U+0A70, U+0A81..0A82 ...)
  \p{Indic_Syllabic_Category: Brahmi_Joining_Number} (Short:
                \p{InSC=BrahmiJoiningNumber}) (20:
                U+11052..11065)
  \p{Indic_Syllabic_Category: Cantillation_Mark} (Short: \p{InSC=
                CantillationMark}) (59: U+0951..0952,
                U+0A51, U+0AFA..0AFC, U+1CD0..1CD2,
                U+1CD4..1CE1, U+1CF4 ...)
  \p{Indic_Syllabic_Category: Consonant} (Short: \p{InSC=Consonant})
                (2195: U+0915..0939, U+0958..095F,
                U+0978..097F, U+0995..09A8,
                U+09AA..09B0, U+09B2 ...)
  \p{Indic_Syllabic_Category: Consonant_Dead} (Short: \p{InSC=
                ConsonantDead}) (12: U+09CE,
                U+0D54..0D56, U+0D7A..0D7F, U+1CF2..1CF3)
  \p{Indic_Syllabic_Category: Consonant_Final} (Short: \p{InSC=
                ConsonantFinal}) (67: U+1930..1931,
                U+1933..1939, U+19C1..19C7,
                U+1A58..1A59, U+1BBE..1BBF, U+1BF0..1BF1
                ...)
  \p{Indic_Syllabic_Category: Consonant_Head_Letter} (Short:
                \p{InSC=ConsonantHeadLetter}) (5:
                U+0F88..0F8C)
  \p{Indic_Syllabic_Category: Consonant_Initial_Postfixed} (Short:
                \p{InSC=ConsonantInitialPostfixed}) (1:
                U+1A5A)
  \p{Indic_Syllabic_Category: Consonant_Killer} (Short: \p{InSC=
                ConsonantKiller}) (2: U+0E4C, U+17CD)
  \p{Indic_Syllabic_Category: Consonant_Medial} (Short: \p{InSC=
                ConsonantMedial}) (31: U+0A75,
                U+0EBC..0EBD, U+103B..103E,
                U+105E..1060, U+1082, U+1A55..1A56 ...)
  \p{Indic_Syllabic_Category: Consonant_Placeholder} (Short:
                \p{InSC=ConsonantPlaceholder}) (22: [\-

\xa0\xd7], U+0980, U+0A72..0A73, U+104B,
U+104E, U+1900 ...)

\p{Indic_Syllabic_Category: Consonant_Preceding_Repha} (Short:
\p{InSC=ConsonantPrecedingRepha}) (3:
U+0D4E, U+11941, U+11D46)

\p{Indic_Syllabic_Category: Consonant_Prefixed} (Short: \p{InSC=
ConsonantPrefixed}) (10: U+111C2..111C3,
U+1193F, U+11A3A, U+11A84..11A89)

\p{Indic_Syllabic_Category: Consonant_Subjoined} (Short: \p{InSC=
ConsonantSubjoined}) (94: U+0F8D..0F97,
U+0F99..0FBC, U+1929..192B, U+1A57,
U+1A5B..1A5E, U+1BA1..1BA3 ...)

\p{Indic_Syllabic_Category: Consonant_Succeeding_Repha} (Short:
\p{InSC=ConsonantSucceedingRepha}) (4:
U+17CC, U+1B03, U+1B81, U+A982)

\p{Indic_Syllabic_Category: Consonant_With_Stacker} (Short:
\p{InSC=ConsonantWithStacker}) (8:
U+0CF1..0CF2, U+1CF5..1CF6,
U+11003..11004, U+11460..11461)

\p{Indic_Syllabic_Category: Gemination_Mark} (Short: \p{InSC=
GeminationMark}) (3: U+0A71, U+11237,
U+11A98)

\p{Indic_Syllabic_Category: Invisible_Stacker} (Short: \p{InSC=
InvisibleStacker}) (12: U+1039, U+17D2,
U+1A60, U+1BAB, U+AAF6, U+10A3F ...)

\p{Indic_Syllabic_Category: Joiner} (Short: \p{InSC=Joiner}) (1:
U+200D)

\p{Indic_Syllabic_Category: Modifying_Letter} (Short: \p{InSC=
ModifyingLetter}) (1: U+0B83)

\p{Indic_Syllabic_Category: Non_Joiner} (Short: \p{InSC=
NonJoiner}) (1: U+200C)

\p{Indic_Syllabic_Category: Nukta} (Short: \p{InSC=Nukta}) (31:
U+093C, U+09BC, U+0A3C, U+0ABC,
U+0AFD..0AFF, U+0B3C ...)

\p{Indic_Syllabic_Category: Number} (Short: \p{InSC=Number}) (491:
[0-9], U+0966..096F, U+09E6..09EF,
U+0A66..0A6F, U+0AE6..0AEF, U+0B66..0B6F
...)

\p{Indic_Syllabic_Category: Number_Joiner} (Short: \p{InSC=
NumberJoiner}) (1: U+1107F)

\p{Indic_Syllabic_Category: Other} (Short: \p{InSC=Other})
                (1_109_572 plus all above-Unicode code
                points: [\x00-\x20!\"#\$\%&\'\(\)*+,.
                \/:;<=>?\@A-Z\[\\\]\^_'a-z\{\|\}~\x7f-
                \x9f\xa1-\xb1\xb4-\xd6\xd8-\xff],
                U+0100..08FF, U+0950, U+0953..0954,
                U+0964..0965, U+0970..0971 ...)
\p{Indic_Syllabic_Category: Pure_Killer} (Short: \p{InSC=
                PureKiller}) (23: U+0D3B..0D3C, U+0E3A,
                U+0E4E, U+0EBA, U+0F84, U+103A ...)
\p{Indic_Syllabic_Category: Register_Shifter} (Short: \p{InSC=
                RegisterShifter}) (2: U+17C9..17CA)
\p{Indic_Syllabic_Category: Syllable_Modifier} (Short: \p{InSC=
                SyllableModifier}) (25: [\xb2-\xb3],
                U+09FE, U+0F35, U+0F37, U+0FC6, U+17CB
                ...)
\p{Indic_Syllabic_Category: Tone_Letter} (Short: \p{InSC=
                ToneLetter}) (7: U+1970..1974, U+AAC0,
                U+AAC2)
\p{Indic_Syllabic_Category: Tone_Mark} (Short: \p{InSC=ToneMark})
                (42: U+0E48..0E4B, U+0EC8..0ECB, U+1037,
                U+1063..1064, U+1069..106D, U+1087..108D
                ...)
\p{Indic_Syllabic_Category: Virama} (Short: \p{InSC=Virama}) (27:
                U+094D, U+09CD, U+0A4D, U+0ACD, U+0B4D,
                U+0BCD ...)
\p{Indic_Syllabic_Category: Visarga} (Short: \p{InSC=Visarga})
                (35: U+0903, U+0983, U+0A03, U+0A83,
                U+0B03, U+0C03 ...)
\p{Indic_Syllabic_Category: Vowel} (Short: \p{InSC=Vowel}) (30:
                U+1963..196D, U+A85E..A861, U+A866,
                U+A922..A92A, U+11150..11154)
\p{Indic_Syllabic_Category: Vowel_Dependent} (Short: \p{InSC=
                VowelDependent}) (683: U+093A..093B,
                U+093E..094C, U+094E..094F,
                U+0955..0957, U+0962..0963, U+09BE..09C4
                ...)
\p{Indic_Syllabic_Category: Vowel_Independent} (Short: \p{InSC=
                VowelIndependent}) (484: U+0904..0914,
                U+0960..0961, U+0972..0977,

```
                    U+0985..098C, U+098F..0990, U+0993..0994
                    ...)
    \p{Inherited}          \p{Script_Extensions=Inherited} (Short:
                    \p{Zinh}) (503)
    \p{Initial_Punctuation} \p{General_Category=Initial_Punctuation}
                    (Short: \p{Pi}) (12)
    \p{InPC: *}           \p{Indic_Positional_Category: *}
    \p{InSC: *}           \p{Indic_Syllabic_Category: *}
    \p{Inscriptional_Pahlavi} \p{Script_Extensions=
                    Inscriptional_Pahlavi} (Short: \p{Phli};
                    NOT \p{Block=Inscriptional_Pahlavi}) (27)
    \p{Inscriptional_Parthian} \p{Script_Extensions=
                    Inscriptional_Parthian} (Short:
                    \p{Prti}; NOT \p{Block=
                    Inscriptional_Parthian}) (30)
  X \p{IPA_Ext}          \p{IPA_Extensions} (= \p{Block=
                    IPA_Extensions}) (96)
  X \p{IPA_Extensions}      \p{Block=IPA_Extensions} (Short:
                    \p{InIPAExt}) (96)
   \p{Is_*}             \p{*} (Any exceptions are individually
                    noted beginning with the word NOT.) If
                    an entry has flag(s) at its beginning,
                    like "D", the "Is_" form has the same
                    flag(s)
   \p{Ital}             \p{Old_Italic} (= \p{Script_Extensions=
                    Old_Italic}) (NOT \p{Block=Old_Italic})
                    (39)
  X \p{Jamo}            \p{Hangul_Jamo} (= \p{Block=Hangul_Jamo})
                    (256)
  X \p{Jamo_Ext_A}         \p{Hangul_Jamo_Extended_A} (= \p{Block=
                    Hangul_Jamo_Extended_A}) (32)
  X \p{Jamo_Ext_B}         \p{Hangul_Jamo_Extended_B} (= \p{Block=
                    Hangul_Jamo_Extended_B}) (80)
   \p{Java}             \p{Javanese} (= \p{Script_Extensions=
                    Javanese}) (NOT \p{Block=Javanese}) (91)
   \p{Javanese}           \p{Script_Extensions=Javanese} (Short:
                    \p{Java}; NOT \p{Block=Javanese}) (91)
   \p{Jg: *}            \p{Joining_Group: *}
   \p{Join_C}            \p{Join_Control} (= \p{Join_Control=Y}) (2)
   \p{Join_C: *}          \p{Join_Control: *}
```

\p{Join_Control}   \p{Join_Control=Y} (Short: \p{JoinC}) (2)
\p{Join_Control: N*}   (Short: \p{JoinC=N}, \P{JoinC}) (1_114_110
                 plus all above-Unicode code points:
                 U+0000..200B, U+200E..infinity)
\p{Join_Control: Y*}   (Short: \p{JoinC=Y}, \p{JoinC}) (2:
                 U+200C..200D)
\p{Joining_Group: African_Feh} (Short: \p{Jg=AfricanFeh}) (1:
                 U+08BB)
\p{Joining_Group: African_Noon} (Short: \p{Jg=AfricanNoon}) (1:
                 U+08BD)
\p{Joining_Group: African_Qaf} (Short: \p{Jg=AfricanQaf}) (2:
                 U+08BC, U+08C4)
\p{Joining_Group: Ain} (Short: \p{Jg=Ain}) (9: U+0639..063A,
                 U+06A0, U+06FC, U+075D..075F, U+08B3,
                 U+08C3)
\p{Joining_Group: Alaph} (Short: \p{Jg=Alaph}) (1: U+0710)
\p{Joining_Group: Alef} (Short: \p{Jg=Alef}) (10: U+0622..0623,
                 U+0625, U+0627, U+0671..0673, U+0675,
                 U+0773..0774)
\p{Joining_Group: Beh} (Short: \p{Jg=Beh}) (27: U+0628,
                 U+062A..062B, U+066E, U+0679..0680,
                 U+0750..0756, U+08A0..08A1 ...)
\p{Joining_Group: Beth} (Short: \p{Jg=Beth}) (2: U+0712, U+072D)
\p{Joining_Group: Burushaski_Yeh_Barree} (Short: \p{Jg=
                 BurushaskiYehBarree}) (2: U+077A..077B)
\p{Joining_Group: Dal} (Short: \p{Jg=Dal}) (15: U+062F..0630,
                 U+0688..0690, U+06EE, U+0759..075A,
                 U+08AE)
\p{Joining_Group: Dalath_Rish} (Short: \p{Jg=DalathRish}) (4:
                 U+0715..0716, U+072A, U+072F)
\p{Joining_Group: E}   (Short: \p{Jg=E}) (1: U+0725)
\p{Joining_Group: Farsi_Yeh} (Short: \p{Jg=FarsiYeh}) (7:
                 U+063D..063F, U+06CC, U+06CE,
                 U+0775..0776)
\p{Joining_Group: Fe}   (Short: \p{Jg=Fe}) (1: U+074F)
\p{Joining_Group: Feh} (Short: \p{Jg=Feh}) (10: U+0641,
                 U+06A1..06A6, U+0760..0761, U+08A4)
\p{Joining_Group: Final_Semkath} (Short: \p{Jg=FinalSemkath}) (1:
                 U+0724)
\p{Joining_Group: Gaf} (Short: \p{Jg=Gaf}) (15: U+063B..063C,

U+06A9, U+06AB, U+06AF..06B4,
U+0762..0764, U+08B0 ...)

\p{Joining_Group: Gamal} (Short: \p{Jg=Gamal}) (3: U+0713..0714,
U+072E)

\p{Joining_Group: Hah} (Short: \p{Jg=Hah}) (21: U+062C..062E,
U+0681..0687, U+06BF, U+0757..0758,
U+076E..076F, U+0772 ...)

\p{Joining_Group: Hamza_On_Heh_Goal} (Short: \p{Jg=
HamzaOnHehGoal}) (1: U+06C3)

\p{Joining_Group: Hanifi_Rohingya_Kinna_Ya} (Short: \p{Jg=
HanifiRohingyaKinnaYa}) (4: U+10D19,
U+10D1E, U+10D20, U+10D23)

\p{Joining_Group: Hanifi_Rohingya_Pa} (Short: \p{Jg=
HanifiRohingyaPa}) (3: U+10D02, U+10D09,
U+10D1C)

\p{Joining_Group: He}   (Short: \p{Jg=He}) (1: U+0717)

\p{Joining_Group: Heh} (Short: \p{Jg=Heh}) (1: U+0647)

\p{Joining_Group: Heh_Goal} (Short: \p{Jg=HehGoal}) (2:
U+06C1..06C2)

\p{Joining_Group: Heth} (Short: \p{Jg=Heth}) (1: U+071A)

\p{Joining_Group: Kaf} (Short: \p{Jg=Kaf}) (6: U+0643,
U+06AC..06AE, U+077F, U+08B4)

\p{Joining_Group: Kaph} (Short: \p{Jg=Kaph}) (1: U+071F)

\p{Joining_Group: Khaph} (Short: \p{Jg=Khaph}) (1: U+074E)

\p{Joining_Group: Knotted_Heh} (Short: \p{Jg=KnottedHeh}) (2:
U+06BE, U+06FF)

\p{Joining_Group: Lam} (Short: \p{Jg=Lam}) (8: U+0644,
U+06B5..06B8, U+076A, U+08A6, U+08C7)

\p{Joining_Group: Lamadh} (Short: \p{Jg=Lamadh}) (1: U+0720)

\p{Joining_Group: Malayalam_Bha} (Short: \p{Jg=MalayalamBha}) (1:
U+0866)

\p{Joining_Group: Malayalam_Ja} (Short: \p{Jg=MalayalamJa}) (1:
U+0861)

\p{Joining_Group: Malayalam_Lla} (Short: \p{Jg=MalayalamLla}) (1:
U+0868)

\p{Joining_Group: Malayalam_Llla} (Short: \p{Jg=MalayalamLlla})
(1: U+0869)

\p{Joining_Group: Malayalam_Nga} (Short: \p{Jg=MalayalamNga}) (1:
U+0860)

\p{Joining_Group: Malayalam_Nna} (Short: \p{Jg=MalayalamNna}) (1:

U+0864)

\p{Joining_Group: Malayalam_Nnna} (Short: \p{Jg=MalayalamNnna})
          (1: U+0865)

\p{Joining_Group: Malayalam_Nya} (Short: \p{Jg=MalayalamNya}) (1:
          U+0862)

\p{Joining_Group: Malayalam_Ra} (Short: \p{Jg=MalayalamRa}) (1:
          U+0867)

\p{Joining_Group: Malayalam_Ssa} (Short: \p{Jg=MalayalamSsa}) (1:
          U+086A)

\p{Joining_Group: Malayalam_Tta} (Short: \p{Jg=MalayalamTta}) (1:
          U+0863)

\p{Joining_Group: Manichaean_Aleph} (Short: \p{Jg=
          ManichaeanAleph}) (1: U+10AC0)

\p{Joining_Group: Manichaean_Ayin} (Short: \p{Jg=ManichaeanAyin})
          (2: U+10AD9..10ADA)

\p{Joining_Group: Manichaean_Beth} (Short: \p{Jg=ManichaeanBeth})
          (2: U+10AC1..10AC2)

\p{Joining_Group: Manichaean_Daleth} (Short: \p{Jg=
          ManichaeanDaleth}) (1: U+10AC5)

\p{Joining_Group: Manichaean_Dhamedh} (Short: \p{Jg=
          ManichaeanDhamedh}) (1: U+10AD4)

\p{Joining_Group: Manichaean_Five} (Short: \p{Jg=ManichaeanFive})
          (1: U+10AEC)

\p{Joining_Group: Manichaean_Gimel} (Short: \p{Jg=
          ManichaeanGimel}) (2: U+10AC3..10AC4)

\p{Joining_Group: Manichaean_Heth} (Short: \p{Jg=ManichaeanHeth})
          (1: U+10ACD)

\p{Joining_Group: Manichaean_Hundred} (Short: \p{Jg=
          ManichaeanHundred}) (1: U+10AEF)

\p{Joining_Group: Manichaean_Kaph} (Short: \p{Jg=ManichaeanKaph})
          (3: U+10AD0..10AD2)

\p{Joining_Group: Manichaean_Lamedh} (Short: \p{Jg=
          ManichaeanLamedh}) (1: U+10AD3)

\p{Joining_Group: Manichaean_Mem} (Short: \p{Jg=ManichaeanMem})
          (1: U+10AD6)

\p{Joining_Group: Manichaean_Nun} (Short: \p{Jg=ManichaeanNun})
          (1: U+10AD7)

\p{Joining_Group: Manichaean_One} (Short: \p{Jg=ManichaeanOne})
          (1: U+10AEB)

\p{Joining_Group: Manichaean_Pe} (Short: \p{Jg=ManichaeanPe}) (2:

U+10ADB..10ADC)
\p{Joining_Group: Manichaean_Qoph} (Short: \p{Jg=ManichaeanQoph})
            (3: U+10ADE..10AE0)
\p{Joining_Group: Manichaean_Resh} (Short: \p{Jg=ManichaeanResh})
            (1: U+10AE1)
\p{Joining_Group: Manichaean_Sadhe} (Short: \p{Jg=
            ManichaeanSadhe}) (1: U+10ADD)
\p{Joining_Group: Manichaean_Samekh} (Short: \p{Jg=
            ManichaeanSamekh}) (1: U+10AD8)
\p{Joining_Group: Manichaean_Taw} (Short: \p{Jg=ManichaeanTaw})
            (1: U+10AE4)
\p{Joining_Group: Manichaean_Ten} (Short: \p{Jg=ManichaeanTen})
            (1: U+10AED)
\p{Joining_Group: Manichaean_Teth} (Short: \p{Jg=ManichaeanTeth})
            (1: U+10ACE)
\p{Joining_Group: Manichaean_Thamedh} (Short: \p{Jg=
            ManichaeanThamedh}) (1: U+10AD5)
\p{Joining_Group: Manichaean_Twenty} (Short: \p{Jg=
            ManichaeanTwenty}) (1: U+10AEE)
\p{Joining_Group: Manichaean_Waw} (Short: \p{Jg=ManichaeanWaw})
            (1: U+10AC7)
\p{Joining_Group: Manichaean_Yodh} (Short: \p{Jg=ManichaeanYodh})
            (1: U+10ACF)
\p{Joining_Group: Manichaean_Zayin} (Short: \p{Jg=
            ManichaeanZayin}) (2: U+10AC9..10ACA)
\p{Joining_Group: Meem} (Short: \p{Jg=Meem}) (4: U+0645,
            U+0765..0766, U+08A7)
\p{Joining_Group: Mim} (Short: \p{Jg=Mim}) (1: U+0721)
\p{Joining_Group: No_Joining_Group} (Short: \p{Jg=NoJoiningGroup})
            (1_113_790 plus all above-Unicode code
            points: U+0000..061F, U+0621, U+0640,
            U+064B..066D, U+0670, U+0674 ...)
\p{Joining_Group: Noon} (Short: \p{Jg=Noon}) (8: U+0646,
            U+06B9..06BC, U+0767..0769)
\p{Joining_Group: Nun} (Short: \p{Jg=Nun}) (1: U+0722)
\p{Joining_Group: Nya} (Short: \p{Jg=Nya}) (1: U+06BD)
\p{Joining_Group: Pe}   (Short: \p{Jg=Pe}) (1: U+0726)
\p{Joining_Group: Qaf} (Short: \p{Jg=Qaf}) (5: U+0642, U+066F,
            U+06A7..06A8, U+08A5)
\p{Joining_Group: Qaph} (Short: \p{Jg=Qaph}) (1: U+0729)

\p{Joining_Group: Reh} (Short: \p{Jg=Reh}) (19: U+0631..0632,
                U+0691..0699, U+06EF, U+075B,
                U+076B..076C, U+0771 ...)
\p{Joining_Group: Reversed_Pe} (Short: \p{Jg=ReversedPe}) (1:
                U+0727)
\p{Joining_Group: Rohingya_Yeh} (Short: \p{Jg=RohingyaYeh}) (1:
                U+08AC)
\p{Joining_Group: Sad} (Short: \p{Jg=Sad}) (6: U+0635..0636,
                U+069D..069E, U+06FB, U+08AF)
\p{Joining_Group: Sadhe} (Short: \p{Jg=Sadhe}) (1: U+0728)
\p{Joining_Group: Seen} (Short: \p{Jg=Seen}) (11: U+0633..0634,
                U+069A..069C, U+06FA, U+075C, U+076D,
                U+0770 ...)
\p{Joining_Group: Semkath} (Short: \p{Jg=Semkath}) (1: U+0723)
\p{Joining_Group: Shin} (Short: \p{Jg=Shin}) (1: U+072B)
\p{Joining_Group: Straight_Waw} (Short: \p{Jg=StraightWaw}) (1:
                U+08B1)
\p{Joining_Group: Swash_Kaf} (Short: \p{Jg=SwashKaf}) (1: U+06AA)
\p{Joining_Group: Syriac_Waw} (Short: \p{Jg=SyriacWaw}) (1: U+0718)
\p{Joining_Group: Tah} (Short: \p{Jg=Tah}) (4: U+0637..0638,
                U+069F, U+08A3)
\p{Joining_Group: Taw} (Short: \p{Jg=Taw}) (1: U+072C)
\p{Joining_Group: Teh_Marbuta} (Short: \p{Jg=TehMarbuta}) (3:
                U+0629, U+06C0, U+06D5)
\p{Joining_Group: Teh_Marbuta_Goal} \p{Joining_Group=
                Hamza_On_Heh_Goal} (1)
\p{Joining_Group: Teth} (Short: \p{Jg=Teth}) (2: U+071B..071C)
\p{Joining_Group: Waw} (Short: \p{Jg=Waw}) (16: U+0624, U+0648,
                U+0676..0677, U+06C4..06CB, U+06CF,
                U+0778..0779 ...)
\p{Joining_Group: Yeh} (Short: \p{Jg=Yeh}) (11: U+0620, U+0626,
                U+0649..064A, U+0678, U+06D0..06D1,
                U+0777 ...)
\p{Joining_Group: Yeh_Barree} (Short: \p{Jg=YehBarree}) (2:
                U+06D2..06D3)
\p{Joining_Group: Yeh_With_Tail} (Short: \p{Jg=YehWithTail}) (1:
                U+06CD)
\p{Joining_Group: Yudh} (Short: \p{Jg=Yudh}) (1: U+071D)
\p{Joining_Group: Yudh_He} (Short: \p{Jg=YudhHe}) (1: U+071E)
\p{Joining_Group: Zain} (Short: \p{Jg=Zain}) (1: U+0719)

\p{Joining_Group: Zhain} (Short: \p{Jg=Zhain}) (1: U+074D)
\p{Joining_Type: C}   \p{Joining_Type=Join_Causing} (4)
\p{Joining_Type: D}   \p{Joining_Type=Dual_Joining} (586)
\p{Joining_Type: Dual_Joining} (Short: \p{Jt=D}) (586: U+0620,
                U+0626, U+0628, U+062A..062E,
                U+0633..063F, U+0641..0647 ...)
\p{Joining_Type: Join_Causing} (Short: \p{Jt=C}) (4: U+0640,
                U+07FA, U+180A, U+200D)
\p{Joining_Type: L}   \p{Joining_Type=Left_Joining} (5)
\p{Joining_Type: Left_Joining} (Short: \p{Jt=L}) (5: U+A872,
                U+10ACD, U+10AD7, U+10D00, U+10FCB)
\p{Joining_Type: Non_Joining} (Short: \p{Jt=U}) (1_111_390 plus
                all above-Unicode code points: [\x00-
                \xac\xae-\xff], U+0100..02FF,
                U+0370..0482, U+048A..0590, U+05BE,
                U+05C0 ...)
\p{Joining_Type: R}   \p{Joining_Type=Right_Joining} (130)
\p{Joining_Type: Right_Joining} (Short: \p{Jt=R}) (130:
                U+0622..0625, U+0627, U+0629,
                U+062F..0632, U+0648, U+0671..0673 ...)
\p{Joining_Type: T}   \p{Joining_Type=Transparent} (1997)
\p{Joining_Type: Transparent} (Short: \p{Jt=T}) (1997: [\xad],
                U+0300..036F, U+0483..0489,
                U+0591..05BD, U+05BF, U+05C1..05C2 ...)
\p{Joining_Type: U}   \p{Joining_Type=Non_Joining} (1_111_390
                plus all above-Unicode code points)
\p{Jt: *}   \p{Joining_Type: *}
\p{Kaithi}   \p{Script_Extensions=Kaithi} (Short:
                \p{Kthi}; NOT \p{Block=Kaithi}) (87)
\p{Kali}   \p{Kayah_Li} (= \p{Script_Extensions=
                Kayah_Li}) (48)
\p{Kana}   \p{Katakana} (= \p{Script_Extensions=
                Katakana}) (NOT \p{Block=Katakana}) (356)
X \p{Kana_Ext_A}   \p{Kana_Extended_A} (= \p{Block=
                Kana_Extended_A}) (48)
X \p{Kana_Extended_A}   \p{Block=Kana_Extended_A} (Short:
                \p{InKanaExtA}) (48)
X \p{Kana_Sup}   \p{Kana_Supplement} (= \p{Block=
                Kana_Supplement}) (256)
X \p{Kana_Supplement}   \p{Block=Kana_Supplement} (Short:

```
              \p{InKanaSup}) (256)
X \p{Kanbun}          \p{Block=Kanbun} (16)
X \p{Kangxi}          \p{Kangxi_Radicals} (= \p{Block=
              Kangxi_Radicals}) (224)
X \p{Kangxi_Radicals}    \p{Block=Kangxi_Radicals} (Short:
              \p{InKangxi}) (224)
 \p{Kannada}          \p{Script_Extensions=Kannada} (Short:
              \p{Knda}; NOT \p{Block=Kannada}) (104)
 \p{Katakana}          \p{Script_Extensions=Katakana} (Short:
              \p{Kana}; NOT \p{Block=Katakana}) (356)
X \p{Katakana_Ext}     \p{Katakana_Phonetic_Extensions} (=
              \p{Block=Katakana_Phonetic_Extensions})
              (16)
X \p{Katakana_Phonetic_Extensions} \p{Block=
              Katakana_Phonetic_Extensions} (Short:
              \p{InKatakanaExt}) (16)
 \p{Kayah_Li}          \p{Script_Extensions=Kayah_Li} (Short:
              \p{Kali}) (48)
 \p{Khar}          \p{Kharoshthi} (= \p{Script_Extensions=
              Kharoshthi}) (NOT \p{Block=Kharoshthi})
              (68)
 \p{Kharoshthi}          \p{Script_Extensions=Kharoshthi} (Short:
              \p{Khar}; NOT \p{Block=Kharoshthi}) (68)
 \p{Khitan_Small_Script} \p{Script_Extensions=Khitan_Small_Script}
              (Short: \p{Kits}; NOT \p{Block=
              Khitan_Small_Script}) (471)
 \p{Khmer}          \p{Script_Extensions=Khmer} (Short:
              \p{Khmr}; NOT \p{Block=Khmer}) (146)
X \p{Khmer_Symbols}     \p{Block=Khmer_Symbols} (32)
 \p{Khmr}          \p{Khmer} (= \p{Script_Extensions=Khmer})
              (NOT \p{Block=Khmer}) (146)
 \p{Khoj}          \p{Khojki} (= \p{Script_Extensions=
              Khojki}) (NOT \p{Block=Khojki}) (82)
 \p{Khojki}          \p{Script_Extensions=Khojki} (Short:
              \p{Khoj}; NOT \p{Block=Khojki}) (82)
 \p{Khudawadi}          \p{Script_Extensions=Khudawadi} (Short:
              \p{Sind}; NOT \p{Block=Khudawadi}) (81)
 \p{Kits}          \p{Khitan_Small_Script} (=
              \p{Script_Extensions=
              Khitan_Small_Script}) (NOT \p{Block=
```

                Khitan_Small_Script}) (471)

\p{Knda}   \p{Kannada} (= \p{Script_Extensions=
                Kannada}) (NOT \p{Block=Kannada}) (104)

\p{Kthi}   \p{Kaithi} (= \p{Script_Extensions=
                Kaithi}) (NOT \p{Block=Kaithi}) (87)

\p{L} \pL   \p{Letter} (= \p{General_Category=Letter})
                (131_241)

X \p{L&}   \p{Cased_Letter} (= \p{General_Category=
                Cased_Letter}) (3977)

X \p{L_}   \p{Cased_Letter} (= \p{General_Category=
                Cased_Letter}) Note the trailing '_'
                matters in spite of loose matching
                rules. (3977)

\p{Lana}   \p{Tai_Tham} (= \p{Script_Extensions=
                Tai_Tham}) (NOT \p{Block=Tai_Tham}) (127)

\p{Lao}   \p{Script_Extensions=Lao} (NOT \p{Block=
                Lao}) (82)

\p{Laoo}   \p{Lao} (= \p{Script_Extensions=Lao}) (NOT
                \p{Block=Lao}) (82)

\p{Latin}   \p{Script_Extensions=Latin} (Short:
                \p{Latn}) (1403)

X \p{Latin_1}   \p{Latin_1_Supplement} (= \p{Block=
                Latin_1_Supplement}) (128)

X \p{Latin_1_Sup}   \p{Latin_1_Supplement} (= \p{Block=
                Latin_1_Supplement}) (128)

X \p{Latin_1_Supplement} \p{Block=Latin_1_Supplement} (Short:
                \p{InLatin1}) (128)

X \p{Latin_Ext_A}   \p{Latin_Extended_A} (= \p{Block=
                Latin_Extended_A}) (128)

X \p{Latin_Ext_Additional} \p{Latin_Extended_Additional} (=
                \p{Block=Latin_Extended_Additional})
                (256)

X \p{Latin_Ext_B}   \p{Latin_Extended_B} (= \p{Block=
                Latin_Extended_B}) (208)

X \p{Latin_Ext_C}   \p{Latin_Extended_C} (= \p{Block=
                Latin_Extended_C}) (32)

X \p{Latin_Ext_D}   \p{Latin_Extended_D} (= \p{Block=
                Latin_Extended_D}) (224)

X \p{Latin_Ext_E}   \p{Latin_Extended_E} (= \p{Block=
                Latin_Extended_E}) (64)

X \p{Latin_Extended_A}   \p{Block=Latin_Extended_A} (Short:
         \p{InLatinExtA}) (128)
X \p{Latin_Extended_Additional} \p{Block=Latin_Extended_Additional}
         (Short: \p{InLatinExtAdditional}) (256)
X \p{Latin_Extended_B}   \p{Block=Latin_Extended_B} (Short:
         \p{InLatinExtB}) (208)
X \p{Latin_Extended_C}   \p{Block=Latin_Extended_C} (Short:
         \p{InLatinExtC}) (32)
X \p{Latin_Extended_D}   \p{Block=Latin_Extended_D} (Short:
         \p{InLatinExtD}) (224)
X \p{Latin_Extended_E}   \p{Block=Latin_Extended_E} (Short:
         \p{InLatinExtE}) (64)
 \p{Latn}   \p{Latin} (= \p{Script_Extensions=Latin})
         (1403)
 \p{Lb: *}   \p{Line_Break: *}
 \p{LC}   \p{Cased_Letter} (= \p{General_Category=
         Cased_Letter}) (3977)
 \p{Lepc}   \p{Lepcha} (= \p{Script_Extensions=
         Lepcha}) (NOT \p{Block=Lepcha}) (74)
 \p{Lepcha}   \p{Script_Extensions=Lepcha} (Short:
         \p{Lepc}; NOT \p{Block=Lepcha}) (74)
 \p{Letter}   \p{General_Category=Letter} (Short: \p{L})
         (131_241)
 \p{Letter_Number}   \p{General_Category=Letter_Number} (Short:
         \p{Nl}) (236)
X \p{Letterlike_Symbols} \p{Block=Letterlike_Symbols} (80)
 \p{Limb}   \p{Limbu} (= \p{Script_Extensions=Limbu})
         (NOT \p{Block=Limbu}) (69)
 \p{Limbu}   \p{Script_Extensions=Limbu} (Short:
         \p{Limb}; NOT \p{Block=Limbu}) (69)
 \p{Lina}   \p{Linear_A} (= \p{Script_Extensions=
         Linear_A}) (NOT \p{Block=Linear_A}) (386)
 \p{Linb}   \p{Linear_B} (= \p{Script_Extensions=
         Linear_B}) (268)
 \p{Line_Break: AI}   \p{Line_Break=Ambiguous} (707)
 \p{Line_Break: AL}   \p{Line_Break=Alphabetic} (21_400)
 \p{Line_Break: Alphabetic} (Short: \p{Lb=AL}) (21_400: [#&*<=>\@A-
         Z\^_'a-z~\xa6\xa9\xac\xae-\xaf\xb5\xc0-
         \xd6\xd8-\xf6\xf8-\xff], U+0100..02C6,
         U+02CE..02CF, U+02D1..02D7, U+02DC,

                        U+02DE ...)
    \p{Line_Break: Ambiguous} (Short: \p{Lb=AI}) (707: [\xa7-\xa8\xaa
                        \xb2-\xb3\xb6-\xba\xbc-\xbe\xd7\xf7],
                        U+02C7, U+02C9..02CB, U+02CD, U+02D0,
                        U+02D8..02DB ...)
    \p{Line_Break: B2}   \p{Line_Break=Break_Both} (3)
    \p{Line_Break: BA}   \p{Line_Break=Break_After} (244)
    \p{Line_Break: BB}   \p{Line_Break=Break_Before} (45)
    \p{Line_Break: BK}   \p{Line_Break=Mandatory_Break} (4)
    \p{Line_Break: Break_After} (Short: \p{Lb=BA}) (244: [\t\|\xad],
                        U+058A, U+05BE, U+0964..0965,
                        U+0E5A..0E5B, U+0F0B ...)
    \p{Line_Break: Break_Before} (Short: \p{Lb=BB}) (45: [\xb4],
                        U+02C8, U+02CC, U+02DF, U+0C77, U+0C84
                        ...)
    \p{Line_Break: Break_Both} (Short: \p{Lb=B2}) (3: U+2014,
                        U+2E3A..2E3B)
    \p{Line_Break: Break_Symbols} (Short: \p{Lb=SY}) (1: [\/])
    \p{Line_Break: Carriage_Return} (Short: \p{Lb=CR}) (1: [\r])
    \p{Line_Break: CB}   \p{Line_Break=Contingent_Break} (1)
    \p{Line_Break: CJ}   \p{Line_Break=
                        Conditional_Japanese_Starter} (58)
    \p{Line_Break: CL}   \p{Line_Break=Close_Punctuation} (91)
    \p{Line_Break: Close_Parenthesis} (Short: \p{Lb=CP}) (2: [\)\]])
    \p{Line_Break: Close_Punctuation} (Short: \p{Lb=CL}) (91: [\}],
                        U+0F3B, U+0F3D, U+169C, U+2046, U+207E
                        ...)
    \p{Line_Break: CM}   \p{Line_Break=Combining_Mark} (2286)
    \p{Line_Break: Combining_Mark} (Short: \p{Lb=CM}) (2286: [^\t\n
                        \cK\f\r\x20-\x7e\x85\xa0-\xff],
                        U+0300..034E, U+0350..035B,
                        U+0363..036F, U+0483..0489, U+0591..05BD
                        ...)
    \p{Line_Break: Complex_Context} (Short: \p{Lb=SA}) (750:
                        U+0E01..0E3A, U+0E40..0E4E,
                        U+0E81..0E82, U+0E84, U+0E86..0E8A,
                        U+0E8C..0EA3 ...)
    \p{Line_Break: Conditional_Japanese_Starter} (Short: \p{Lb=CJ})
                        (58: U+3041, U+3043, U+3045, U+3047,
                        U+3049, U+3063 ...)

\p{Line_Break: Contingent_Break} (Short: \p{Lb=CB}) (1: U+FFFC)
\p{Line_Break: CP}   \p{Line_Break=Close_Parenthesis} (2)
\p{Line_Break: CR}   \p{Line_Break=Carriage_Return} (1)
\p{Line_Break: E_Base} (Short: \p{Lb=EB}) (122: U+261D, U+26F9,
                U+270A..270D, U+1F385, U+1F3C2..1F3C4,
                U+1F3C7 ...)
\p{Line_Break: E_Modifier} (Short: \p{Lb=EM}) (5: U+1F3FB..1F3FF)
\p{Line_Break: EB}   \p{Line_Break=E_Base} (122)
\p{Line_Break: EM}   \p{Line_Break=E_Modifier} (5)
\p{Line_Break: EX}   \p{Line_Break=Exclamation} (37)
\p{Line_Break: Exclamation} (Short: \p{Lb=EX}) (37: [!?], U+05C6,
                U+061B, U+061E..061F, U+06D4, U+07F9 ...)
\p{Line_Break: GL}   \p{Line_Break=Glue} (26)
\p{Line_Break: Glue}   (Short: \p{Lb=GL}) (26: [\xa0], U+034F,
                U+035C..0362, U+0F08, U+0F0C, U+0F12 ...)
\p{Line_Break: H2}   (Short: \p{Lb=H2}) (399: U+AC00, U+AC1C,
                U+AC38, U+AC54, U+AC70, U+AC8C ...)
\p{Line_Break: H3}   (Short: \p{Lb=H3}) (10_773: U+AC01..AC1B,
                U+AC1D..AC37, U+AC39..AC53,
                U+AC55..AC6F, U+AC71..AC8B, U+AC8D..ACA7
                ...)
\p{Line_Break: Hebrew_Letter} (Short: \p{Lb=HL}) (75:
                U+05D0..05EA, U+05EF..05F2, U+FB1D,
                U+FB1F..FB28, U+FB2A..FB36, U+FB38..FB3C
                ...)
\p{Line_Break: HL}   \p{Line_Break=Hebrew_Letter} (75)
\p{Line_Break: HY}   \p{Line_Break=Hyphen} (1)
\p{Line_Break: Hyphen} (Short: \p{Lb=HY}) (1: [\-])
\p{Line_Break: ID}   \p{Line_Break=Ideographic} (172_462)
\p{Line_Break: Ideographic} (Short: \p{Lb=ID}) (172_462:
                U+231A..231B, U+23F0..23F3,
                U+2600..2603, U+2614..2615, U+2618,
                U+261A..261C ...)
\p{Line_Break: IN}   \p{Line_Break=Inseparable} (6)
\p{Line_Break: Infix_Numeric} (Short: \p{Lb=IS}) (13: [,.:;],
                U+037E, U+0589, U+060C..060D, U+07F8,
                U+2044 ...)
\p{Line_Break: Inseparable} (Short: \p{Lb=IN}) (6: U+2024..2026,
                U+22EF, U+FE19, U+10AF6)
\p{Line_Break: Inseperable} \p{Line_Break=Inseparable} (6)

\p{Line_Break: IS}   \p{Line_Break=Infix_Numeric} (13)

\p{Line_Break: JL}   (Short: \p{Lb=JL}) (125: U+1100..115F,
                U+A960..A97C)

\p{Line_Break: JT}   (Short: \p{Lb=JT}) (137: U+11A8..11FF,
                U+D7CB..D7FB)

\p{Line_Break: JV}   (Short: \p{Lb=JV}) (95: U+1160..11A7,
                U+D7B0..D7C6)

\p{Line_Break: LF}   \p{Line_Break=Line_Feed} (1)

\p{Line_Break: Line_Feed} (Short: \p{Lb=LF}) (1: [\n])

\p{Line_Break: Mandatory_Break} (Short: \p{Lb=BK}) (4: [\cK\f],
                U+2028..2029)

\p{Line_Break: Next_Line} (Short: \p{Lb=NL}) (1: [\x85])

\p{Line_Break: NL}   \p{Line_Break=Next_Line} (1)

\p{Line_Break: Nonstarter} (Short: \p{Lb=NS}) (33: U+17D6,
                U+203C..203D, U+2047..2049, U+3005,
                U+301C, U+303B..303C ...)

\p{Line_Break: NS}   \p{Line_Break=Nonstarter} (33)

\p{Line_Break: NU}   \p{Line_Break=Numeric} (642)

\p{Line_Break: Numeric} (Short: \p{Lb=NU}) (642: [0-9],
                U+0660..0669, U+066B..066C,
                U+06F0..06F9, U+07C0..07C9, U+0966..096F
                ...)

\p{Line_Break: OP}   \p{Line_Break=Open_Punctuation} (88)

\p{Line_Break: Open_Punctuation} (Short: \p{Lb=OP}) (88: [\(\[\{
                \xa1\xbf], U+0F3A, U+0F3C, U+169B,
                U+201A, U+201E ...)

\p{Line_Break: PO}   \p{Line_Break=Postfix_Numeric} (36)

\p{Line_Break: Postfix_Numeric} (Short: \p{Lb=PO}) (36: [\%\xa2
                \xb0], U+0609..060B, U+066A,
                U+09F2..09F3, U+09F9, U+0D79 ...)

\p{Line_Break: PR}   \p{Line_Break=Prefix_Numeric} (68)

\p{Line_Break: Prefix_Numeric} (Short: \p{Lb=PR}) (68: [\$+\\\xa3-
                \xa5\xb1], U+058F, U+07FE..07FF, U+09FB,
                U+0AF1, U+0BF9 ...)

\p{Line_Break: QU}   \p{Line_Break=Quotation} (39)

\p{Line_Break: Quotation} (Short: \p{Lb=QU}) (39: [\"\'\xab\xbb],
                U+2018..2019, U+201B..201D, U+201F,
                U+2039..203A, U+275B..2760 ...)

\p{Line_Break: Regional_Indicator} (Short: \p{Lb=RI}) (26:
                U+1F1E6..1F1FF)

```
  \p{Line_Break: RI}     \p{Line_Break=Regional_Indicator} (26)
  \p{Line_Break: SA}     \p{Line_Break=Complex_Context} (750)
D \p{Line_Break: SG}     \p{Line_Break=Surrogate} (2048)
  \p{Line_Break: SP}     \p{Line_Break=Space} (1)
  \p{Line_Break: Space}  (Short: \p{Lb=SP}) (1: [\x20])
D \p{Line_Break: Surrogate} Surrogates should never appear in well-
                   formed text, and therefore shouldn't be
                   the basis for line breaking (Short:
                   \p{Lb=SG}) (2048: U+D800..DFFF)
  \p{Line_Break: SY}     \p{Line_Break=Break_Symbols} (1)
  \p{Line_Break: Unknown} (Short: \p{Lb=XX}) (901_256 plus all
                   above-Unicode code points: U+0378..0379,
                   U+0380..0383, U+038B, U+038D, U+03A2,
                   U+0530 ...)
  \p{Line_Break: WJ}     \p{Line_Break=Word_Joiner} (2)
  \p{Line_Break: Word_Joiner} (Short: \p{Lb=WJ}) (2: U+2060, U+FEFF)
  \p{Line_Break: XX}     \p{Line_Break=Unknown} (901_256 plus all
                   above-Unicode code points)
  \p{Line_Break: ZW}     \p{Line_Break=ZWSpace} (1)
  \p{Line_Break: ZWJ}    (Short: \p{Lb=ZWJ}) (1: U+200D)
  \p{Line_Break: ZWSpace} (Short: \p{Lb=ZW}) (1: U+200B)
  \p{Line_Separator}     \p{General_Category=Line_Separator}
                   (Short: \p{Zl}) (1)
  \p{Linear_A}           \p{Script_Extensions=Linear_A} (Short:
                   \p{Lina}; NOT \p{Block=Linear_A}) (386)
  \p{Linear_B}           \p{Script_Extensions=Linear_B} (Short:
                   \p{Linb}) (268)
X \p{Linear_B_Ideograms} \p{Block=Linear_B_Ideograms} (128)
X \p{Linear_B_Syllabary} \p{Block=Linear_B_Syllabary} (128)
  \p{Lisu}               \p{Script_Extensions=Lisu} (NOT \p{Block=
                   Lisu}) (49)
X \p{Lisu_Sup}           \p{Lisu_Supplement} (= \p{Block=
                   Lisu_Supplement}) (16)
X \p{Lisu_Supplement}    \p{Block=Lisu_Supplement} (Short:
                   \p{InLisuSup}) (16)
  \p{Ll}                 \p{Lowercase_Letter} (=
                   \p{General_Category=Lowercase_Letter})
                   (/i= General_Category=Cased_Letter)
                   (2155)
  \p{Lm}                 \p{Modifier_Letter} (=
```

```
                 \p{General_Category=Modifier_Letter})
                 (260)
  \p{Lo}            \p{Other_Letter} (= \p{General_Category=
                 Other_Letter}) (127_004)
  \p{LOE}            \p{Logical_Order_Exception} (=
                 \p{Logical_Order_Exception=Y}) (19)
  \p{LOE: *}          \p{Logical_Order_Exception: *}
  \p{Logical_Order_Exception} \p{Logical_Order_Exception=Y} (Short:
                 \p{LOE}) (19)
  \p{Logical_Order_Exception: N*} (Short: \p{LOE=N}, \P{LOE})
                 (1_114_093 plus all above-Unicode code
                 points: U+0000..0E3F, U+0E45..0EBF,
                 U+0EC5..19B4, U+19B8..19B9,
                 U+19BB..AAB4, U+AAB7..AAB8 ...)
  \p{Logical_Order_Exception: Y*} (Short: \p{LOE=Y}, \p{LOE}) (19:
                 U+0E40..0E44, U+0EC0..0EC4,
                 U+19B5..19B7, U+19BA, U+AAB5..AAB6,
                 U+AAB9 ...)
 X \p{Low_Surrogates}    \p{Block=Low_Surrogates} (1024)
  \p{Lower}            \p{XPosixLower} (= \p{Lowercase=Y}) (/i=
                 Cased=Yes) (2344)
  \p{Lower: *}         \p{Lowercase: *}
  \p{Lowercase}         \p{XPosixLower} (= \p{Lowercase=Y}) (/i=
                 Cased=Yes) (2344)
  \p{Lowercase: N*}     (Short: \p{Lower=N}, \P{Lower}; /i= Cased=
                 No) (1_111_768 plus all above-Unicode
                 code points: [\x00-\x20!\"#\$\%&\'
                 \(\)*+,\-.\/0-9:;<=>?\@A-Z\[\\\]\^_‘\{
                 \|\}~\x7f-\xa9\xab-\xb4\xb6-\xb9\xbb-
                 \xde\xf7], U+0100, U+0102, U+0104,
                 U+0106, U+0108 ...)
  \p{Lowercase: Y*}     (Short: \p{Lower=Y}, \p{Lower}; /i= Cased=
                 Yes) (2344: [a-z\xaa\xb5\xba\xdf-\xf6
                 \xf8-\xff], U+0101, U+0103, U+0105,
                 U+0107, U+0109 ...)
  \p{Lowercase_Letter}   \p{General_Category=Lowercase_Letter}
                 (Short: \p{Ll}; /i= General_Category=
                 Cased_Letter) (2155)
  \p{Lt}            \p{Titlecase_Letter} (=
                 \p{General_Category=Titlecase_Letter})
```

	(/i= General_Category=Cased_Letter) (31)
\p{Lu}	\p{Uppercase_Letter} (= \p{General_Category=Uppercase_Letter}) (/i= General_Category=Cased_Letter) (1791)
\p{Lyci}	\p{Lycian} (= \p{Script_Extensions= Lycian}) (NOT \p{Block=Lycian}) (29)
\p{Lycian}	\p{Script_Extensions=Lycian} (Short: \p{Lyci}; NOT \p{Block=Lycian}) (29)
\p{Lydi}	\p{Lydian} (= \p{Script_Extensions= Lydian}) (NOT \p{Block=Lydian}) (27)
\p{Lydian}	\p{Script_Extensions=Lydian} (Short: \p{Lydi}; NOT \p{Block=Lydian}) (27)
\p{M} \pM	\p{Mark} (= \p{General_Category=Mark}) (2295)
\p{Mahajani}	\p{Script_Extensions=Mahajani} (Short: \p{Mahj}; NOT \p{Block=Mahajani}) (61)
\p{Mahj}	\p{Mahajani} (= \p{Script_Extensions= Mahajani}) (NOT \p{Block=Mahajani}) (61)
X \p{Mahjong}	\p{Mahjong_Tiles} (= \p{Block= Mahjong_Tiles}) (48)
X \p{Mahjong_Tiles}	\p{Block=Mahjong_Tiles} (Short: \p{InMahjong}) (48)
\p{Maka}	\p{Makasar} (= \p{Script_Extensions= Makasar}) (NOT \p{Block=Makasar}) (25)
\p{Makasar}	\p{Script_Extensions=Makasar} (Short: \p{Maka}; NOT \p{Block=Makasar}) (25)
\p{Malayalam}	\p{Script_Extensions=Malayalam} (Short: \p{Mlym}; NOT \p{Block=Malayalam}) (126)
\p{Mand}	\p{Mandaic} (= \p{Script_Extensions= Mandaic}) (NOT \p{Block=Mandaic}) (30)
\p{Mandaic}	\p{Script_Extensions=Mandaic} (Short: \p{Mand}; NOT \p{Block=Mandaic}) (30)
\p{Mani}	\p{Manichaean} (= \p{Script_Extensions= Manichaean}) (NOT \p{Block=Manichaean}) (52)
\p{Manichaean}	\p{Script_Extensions=Manichaean} (Short: \p{Mani}; NOT \p{Block=Manichaean}) (52)
\p{Marc}	\p{Marchen} (= \p{Script_Extensions= Marchen}) (NOT \p{Block=Marchen}) (68)

\p{Marchen}   \p{Script_Extensions=Marchen} (Short:
                     \p{Marc}; NOT \p{Block=Marchen}) (68)
\p{Mark}   \p{General_Category=Mark} (Short: \p{M})
                     (2295)
\p{Masaram_Gondi}   \p{Script_Extensions=Masaram_Gondi}
                     (Short: \p{Gonm}; NOT \p{Block=
                     Masaram_Gondi}) (77)
\p{Math}   \p{Math=Y} (2310)
\p{Math: N*}   (Single: \P{Math}) (1_111_802 plus all
                     above-Unicode code points: [\x00-\x20!
                     \"#\$\%&\'\(\)*,\-.\/0-9:;?\@A-Z
                     \[\\\]_'a-z\{\}\x7f-\xab\xad-\xb0\xb2-
                     \xd6\xd8-\xf6\xf8-\xff], U+0100..03CF,
                     U+03D3..03D4, U+03D6..03EF,
                     U+03F2..03F3, U+03F7..0605 ...)
\p{Math: Y*}   (Single: \p{Math}) (2310: [+<=>\^\|~\xac
                     \xb1\xd7\xf7], U+03D0..03D2, U+03D5,
                     U+03F0..03F1, U+03F4..03F6, U+0606..0608
                     ...)
X \p{Math_Alphanum}   \p{Mathematical_Alphanumeric_Symbols} (=
                     \p{Block=
                     Mathematical_Alphanumeric_Symbols})
                     (1024)
X \p{Math_Operators}   \p{Mathematical_Operators} (= \p{Block=
                     Mathematical_Operators}) (256)
\p{Math_Symbol}   \p{General_Category=Math_Symbol} (Short:
                     \p{Sm}) (948)
X \p{Mathematical_Alphanumeric_Symbols} \p{Block=
                     Mathematical_Alphanumeric_Symbols}
                     (Short: \p{InMathAlphanum}) (1024)
X \p{Mathematical_Operators} \p{Block=Mathematical_Operators}
                     (Short: \p{InMathOperators}) (256)
X \p{Mayan_Numerals}   \p{Block=Mayan_Numerals} (32)
\p{Mc}   \p{Spacing_Mark} (= \p{General_Category=
                     Spacing_Mark}) (443)
\p{Me}   \p{Enclosing_Mark} (= \p{General_Category=
                     Enclosing_Mark}) (13)
\p{Medefaidrin}   \p{Script_Extensions=Medefaidrin} (Short:
                     \p{Medf}; NOT \p{Block=Medefaidrin}) (91)
\p{Medf}   \p{Medefaidrin} (= \p{Script_Extensions=

                    Medefaidrin}) (NOT \p{Block=
                    Medefaidrin}) (91)
 \p{Meetei_Mayek}   \p{Script_Extensions=Meetei_Mayek} (Short:
                    \p{Mtei}; NOT \p{Block=Meetei_Mayek})
                    (79)
 X \p{Meetei_Mayek_Ext}   \p{Meetei_Mayek_Extensions} (= \p{Block=
                    Meetei_Mayek_Extensions}) (32)
 X \p{Meetei_Mayek_Extensions} \p{Block=Meetei_Mayek_Extensions}
                    (Short: \p{InMeeteiMayekExt}) (32)
 \p{Mend}   \p{Mende_Kikakui} (= \p{Script_Extensions=
                    Mende_Kikakui}) (NOT \p{Block=
                    Mende_Kikakui}) (213)
 \p{Mende_Kikakui}   \p{Script_Extensions=Mende_Kikakui}
                    (Short: \p{Mend}; NOT \p{Block=
                    Mende_Kikakui}) (213)
 \p{Merc}   \p{Meroitic_Cursive} (=
                    \p{Script_Extensions=Meroitic_Cursive})
                    (NOT \p{Block=Meroitic_Cursive}) (90)
 \p{Mero}   \p{Meroitic_Hieroglyphs} (=
                    \p{Script_Extensions=
                    Meroitic_Hieroglyphs}) (32)
 \p{Meroitic_Cursive}   \p{Script_Extensions=Meroitic_Cursive}
                    (Short: \p{Merc}; NOT \p{Block=
                    Meroitic_Cursive}) (90)
 \p{Meroitic_Hieroglyphs} \p{Script_Extensions=
                    Meroitic_Hieroglyphs} (Short: \p{Mero})
                    (32)
 \p{Miao}   \p{Script_Extensions=Miao} (NOT \p{Block=
                    Miao}) (149)
 X \p{Misc_Arrows}   \p{Miscellaneous_Symbols_And_Arrows} (=
                    \p{Block=
                    Miscellaneous_Symbols_And_Arrows}) (256)
 X \p{Misc_Math_Symbols_A} \p{Miscellaneous_Mathematical_Symbols_A}
                    (= \p{Block=
                    Miscellaneous_Mathematical_Symbols_A})
                    (48)
 X \p{Misc_Math_Symbols_B} \p{Miscellaneous_Mathematical_Symbols_B}
                    (= \p{Block=
                    Miscellaneous_Mathematical_Symbols_B})
                    (128)

X \p{Misc_Pictographs}   \p{Miscellaneous_Symbols_And_Pictographs}
                (= \p{Block=
                Miscellaneous_Symbols_And_Pictographs})
                (768)
X \p{Misc_Symbols}   \p{Miscellaneous_Symbols} (= \p{Block=
                Miscellaneous_Symbols}) (256)
X \p{Misc_Technical}   \p{Miscellaneous_Technical} (= \p{Block=
                Miscellaneous_Technical}) (256)
X \p{Miscellaneous_Mathematical_Symbols_A} \p{Block=
                Miscellaneous_Mathematical_Symbols_A}
                (Short: \p{InMiscMathSymbolsA}) (48)
X \p{Miscellaneous_Mathematical_Symbols_B} \p{Block=
                Miscellaneous_Mathematical_Symbols_B}
                (Short: \p{InMiscMathSymbolsB}) (128)
X \p{Miscellaneous_Symbols} \p{Block=Miscellaneous_Symbols} (Short:
                \p{InMiscSymbols}) (256)
X \p{Miscellaneous_Symbols_And_Arrows} \p{Block=
                Miscellaneous_Symbols_And_Arrows}
                (Short: \p{InMiscArrows}) (256)
X \p{Miscellaneous_Symbols_And_Pictographs} \p{Block=
                Miscellaneous_Symbols_And_Pictographs}
                (Short: \p{InMiscPictographs}) (768)
X \p{Miscellaneous_Technical} \p{Block=Miscellaneous_Technical}
                (Short: \p{InMiscTechnical}) (256)
 \p{Mlym}   \p{Malayalam} (= \p{Script_Extensions=
                Malayalam}) (NOT \p{Block=Malayalam})
                (126)
 \p{Mn}   \p{Nonspacing_Mark} (=
                \p{General_Category=Nonspacing_Mark})
                (1839)
 \p{Modi}   \p{Script_Extensions=Modi} (NOT \p{Block=
                Modi}) (89)
 \p{Modifier_Letter}   \p{General_Category=Modifier_Letter}
                (Short: \p{Lm}) (260)
X \p{Modifier_Letters}   \p{Spacing_Modifier_Letters} (= \p{Block=
                Spacing_Modifier_Letters}) (80)
 \p{Modifier_Symbol}   \p{General_Category=Modifier_Symbol}
                (Short: \p{Sk}) (123)
X \p{Modifier_Tone_Letters} \p{Block=Modifier_Tone_Letters} (32)
 \p{Mong}   \p{Mongolian} (= \p{Script_Extensions=

```
              Mongolian}) (NOT \p{Block=Mongolian})
              (171)
 \p{Mongolian}       \p{Script_Extensions=Mongolian} (Short:
              \p{Mong}; NOT \p{Block=Mongolian}) (171)
X \p{Mongolian_Sup}    \p{Mongolian_Supplement} (= \p{Block=
              Mongolian_Supplement}) (32)
X \p{Mongolian_Supplement} \p{Block=Mongolian_Supplement} (Short:
              \p{InMongolianSup}) (32)
 \p{Mro}          \p{Script_Extensions=Mro} (NOT \p{Block=
              Mro}) (43)
 \p{Mroo}          \p{Mro} (= \p{Script_Extensions=Mro}) (NOT
              \p{Block=Mro}) (43)
 \p{Mtei}          \p{Meetei_Mayek} (= \p{Script_Extensions=
              Meetei_Mayek}) (NOT \p{Block=
              Meetei_Mayek}) (79)
 \p{Mult}          \p{Multani} (= \p{Script_Extensions=
              Multani}) (NOT \p{Block=Multani}) (48)
 \p{Multani}        \p{Script_Extensions=Multani} (Short:
              \p{Mult}; NOT \p{Block=Multani}) (48)
X \p{Music}         \p{Musical_Symbols} (= \p{Block=
              Musical_Symbols}) (256)
X \p{Musical_Symbols}   \p{Block=Musical_Symbols} (Short:
              \p{InMusic}) (256)
 \p{Myanmar}        \p{Script_Extensions=Myanmar} (Short:
              \p{Mymr}; NOT \p{Block=Myanmar}) (224)
X \p{Myanmar_Ext_A}    \p{Myanmar_Extended_A} (= \p{Block=
              Myanmar_Extended_A}) (32)
X \p{Myanmar_Ext_B}    \p{Myanmar_Extended_B} (= \p{Block=
              Myanmar_Extended_B}) (32)
X \p{Myanmar_Extended_A}  \p{Block=Myanmar_Extended_A} (Short:
              \p{InMyanmarExtA}) (32)
X \p{Myanmar_Extended_B}  \p{Block=Myanmar_Extended_B} (Short:
              \p{InMyanmarExtB}) (32)
 \p{Mymr}          \p{Myanmar} (= \p{Script_Extensions=
              Myanmar}) (NOT \p{Block=Myanmar}) (224)
 \p{N} \pN         \p{Number} (= \p{General_Category=Number})
              (1781)
 \p{Na=*}         \p{Name=*}
 \p{Nabataean}       \p{Script_Extensions=Nabataean} (Short:
              \p{Nbat}; NOT \p{Block=Nabataean}) (40)
```

```
  \p{Name=*}          Combination of Name and Name_Alias
                      properties; has special loose matching
                      rules, for which see Unicode UAX #44
  \p{Nand}            \p{Nandinagari} (= \p{Script_Extensions=
                      Nandinagari}) (NOT \p{Block=
                      Nandinagari}) (86)
  \p{Nandinagari}     \p{Script_Extensions=Nandinagari} (Short:
                      \p{Nand}; NOT \p{Block=Nandinagari}) (86)
  \p{Narb}            \p{Old_North_Arabian} (=
                      \p{Script_Extensions=Old_North_Arabian})
                      (32)
X \p{NB}              \p{No_Block} (= \p{Block=No_Block})
                      (826_640 plus all above-Unicode code
                      points)
  \p{Nbat}            \p{Nabataean} (= \p{Script_Extensions=
                      Nabataean}) (NOT \p{Block=Nabataean})
                      (40)
  \p{NChar}           \p{Noncharacter_Code_Point} (=
                      \p{Noncharacter_Code_Point=Y}) (66)
  \p{NChar: *}        \p{Noncharacter_Code_Point: *}
  \p{Nd}              \p{XPosixDigit} (= \p{General_Category=
                      Decimal_Number}) (650)
  \p{New_Tai_Lue}     \p{Script_Extensions=New_Tai_Lue} (Short:
                      \p{Talu}; NOT \p{Block=New_Tai_Lue}) (83)
  \p{Newa}            \p{Script_Extensions=Newa} (NOT \p{Block=
                      Newa}) (97)
  \p{NFC_QC: *}       \p{NFC_Quick_Check: *}
  \p{NFC_Quick_Check: M}  \p{NFC_Quick_Check=Maybe} (111)
  \p{NFC_Quick_Check: Maybe} (Short: \p{NFCQC=M}) (111:
                      U+0300..0304, U+0306..030C, U+030F,
                      U+0311, U+0313..0314, U+031B ...)
  \p{NFC_Quick_Check: N}  \p{NFC_Quick_Check=No} (NOT
                      \P{NFC_Quick_Check} NOR \P{NFC_QC})
                      (1120)
  \p{NFC_Quick_Check: No} (Short: \p{NFCQC=N}; NOT
                      \P{NFC_Quick_Check} NOR \P{NFC_QC})
                      (1120: U+0340..0341, U+0343..0344,
                      U+0374, U+037E, U+0387, U+0958..095F ...)
  \p{NFC_Quick_Check: Y}  \p{NFC_Quick_Check=Yes} (NOT
                      \p{NFC_Quick_Check} NOR \p{NFC_QC})
```

                    (1_112_881 plus all above-Unicode code
                    points)
\p{NFC_Quick_Check: Yes} (Short: \p{NFCQC=Y}; NOT
                    \p{NFC_Quick_Check} NOR \p{NFC_QC})
                    (1_112_881 plus all above-Unicode code
                    points: U+0000..02FF, U+0305,
                    U+030D..030E, U+0310, U+0312,
                    U+0315..031A ...)
\p{NFD_QC: *}   \p{NFD_Quick_Check: *}
\p{NFD_Quick_Check: N} \p{NFD_Quick_Check=No} (NOT
                    \P{NFD_Quick_Check} NOR \P{NFD_QC})
                    (13_233)
\p{NFD_Quick_Check: No} (Short: \p{NFDQC=N}; NOT
                    \P{NFD_Quick_Check} NOR \P{NFD_QC})
                    (13_233: [\xc0-\xc5\xc7-\xcf\xd1-\xd6
                    \xd9-\xdd\xe0-\xe5\xe7-\xef\xf1-\xf6
                    \xf9-\xfd\xff], U+0100..010F,
                    U+0112..0125, U+0128..0130,
                    U+0134..0137, U+0139..013E ...)
\p{NFD_Quick_Check: Y} \p{NFD_Quick_Check=Yes} (NOT
                    \p{NFD_Quick_Check} NOR \p{NFD_QC})
                    (1_100_879 plus all above-Unicode code
                    points)
\p{NFD_Quick_Check: Yes} (Short: \p{NFDQC=Y}; NOT
                    \p{NFD_Quick_Check} NOR \p{NFD_QC})
                    (1_100_879 plus all above-Unicode code
                    points: [\x00-\xbf\xc6\xd0\xd7-\xd8\xde-
                    \xdf\xe6\xf0\xf7-\xf8\xfe],
                    U+0110..0111, U+0126..0127,
                    U+0131..0133, U+0138, U+013F..0142 ...)
\p{NFKC_QC: *}   \p{NFKC_Quick_Check: *}
\p{NFKC_Quick_Check: M} \p{NFKC_Quick_Check=Maybe} (111)
\p{NFKC_Quick_Check: Maybe} (Short: \p{NFKCQC=M}) (111:
                    U+0300..0304, U+0306..030C, U+030F,
                    U+0311, U+0313..0314, U+031B ...)
\p{NFKC_Quick_Check: N} \p{NFKC_Quick_Check=No} (NOT
                    \P{NFKC_Quick_Check} NOR \P{NFKC_QC})
                    (4807)
\p{NFKC_Quick_Check: No} (Short: \p{NFKCQC=N}; NOT
                    \P{NFKC_Quick_Check} NOR \P{NFKC_QC})

                (4807: [\xa0\xa8\xaa\xaf\xb2-\xb5\xb8-
                \xba\xbc-\xbe], U+0132..0133,
                U+013F..0140, U+0149, U+017F,
                U+01C4..01CC ...)
\p{NFKC_Quick_Check: Y} \p{NFKC_Quick_Check=Yes} (NOT
                \p{NFKC_Quick_Check} NOR \p{NFKC_QC})
                (1_109_194 plus all above-Unicode code
                points)
\p{NFKC_Quick_Check: Yes} (Short: \p{NFKCQC=Y}; NOT
                \p{NFKC_Quick_Check} NOR \p{NFKC_QC})
                (1_109_194 plus all above-Unicode code
                points: [\x00-\x9f\xa1-\xa7\xa9\xab-
                \xae\xb0-\xb1\xb6-\xb7\xbb\xbf-\xff],
                U+0100..0131, U+0134..013E,
                U+0141..0148, U+014A..017E, U+0180..01C3
                ...)
\p{NFKD_QC: *}   \p{NFKD_Quick_Check: *}
\p{NFKD_Quick_Check: N} \p{NFKD_Quick_Check=No} (NOT
                \P{NFKD_Quick_Check} NOR \P{NFKD_QC})
                (16_908)
\p{NFKD_Quick_Check: No} (Short: \p{NFKDQC=N}; NOT
                \P{NFKD_Quick_Check} NOR \P{NFKD_QC})
                (16_908: [\xa0\xa8\xaa\xaf\xb2-\xb5\xb8-
                \xba\xbc-\xbe\xc0-\xc5\xc7-\xcf\xd1-
                \xd6\xd9-\xdd\xe0-\xe5\xe7-\xef\xf1-
                \xf6\xf9-\xfd\xff], U+0100..010F,
                U+0112..0125, U+0128..0130,
                U+0132..0137, U+0139..0140 ...)
\p{NFKD_Quick_Check: Y} \p{NFKD_Quick_Check=Yes} (NOT
                \p{NFKD_Quick_Check} NOR \p{NFKD_QC})
                (1_097_204 plus all above-Unicode code
                points)
\p{NFKD_Quick_Check: Yes} (Short: \p{NFKDQC=Y}; NOT
                \p{NFKD_Quick_Check} NOR \p{NFKD_QC})
                (1_097_204 plus all above-Unicode code
                points: [\x00-\x9f\xa1-\xa7\xa9\xab-
                \xae\xb0-\xb1\xb6-\xb7\xbb\xbf\xc6\xd0
                \xd7-\xd8\xde-\xdf\xe6\xf0\xf7-\xf8
                \xfe], U+0110..0111, U+0126..0127,
                U+0131, U+0138, U+0141..0142 ...)

```
   \p{Nko}              \p{Script_Extensions=Nko} (NOT \p{Block=
                        NKo}) (62)
   \p{Nkoo}             \p{Nko} (= \p{Script_Extensions=Nko}) (NOT
                        \p{Block=NKo}) (62)
   \p{Nl}               \p{Letter_Number} (= \p{General_Category=
                        Letter_Number}) (236)
   \p{No}               \p{Other_Number} (= \p{General_Category=
                        Other_Number}) (895)
 X \p{No_Block}         \p{Block=No_Block} (Short: \p{InNB})
                        (826_640 plus all above-Unicode code
                        points)
   \p{Noncharacter_Code_Point} \p{Noncharacter_Code_Point=Y} (Short:
                        \p{NChar}) (66)
   \p{Noncharacter_Code_Point: N*} (Short: \p{NChar=N}, \P{NChar})
                        (1_114_046 plus all above-Unicode code
                        points: U+0000..FDCF, U+FDF0..FFFD,
                        U+10000..1FFFD, U+20000..2FFFD,
                        U+30000..3FFFD, U+40000..4FFFD ...)
   \p{Noncharacter_Code_Point: Y*} (Short: \p{NChar=Y}, \p{NChar})
                        (66: U+FDD0..FDEF, U+FFFE..FFFF,
                        U+1FFFE..1FFFF, U+2FFFE..2FFFF,
                        U+3FFFE..3FFFF, U+4FFFE..4FFFF ...)
   \p{Nonspacing_Mark}  \p{General_Category=Nonspacing_Mark}
                        (Short: \p{Mn}) (1839)
   \p{Nshu}             \p{Nushu} (= \p{Script_Extensions=Nushu})
                        (NOT \p{Block=Nushu}) (397)
   \p{Nt: *}            \p{Numeric_Type: *}
   \p{Number}           \p{General_Category=Number} (Short: \p{N})
                        (1781)
 X \p{Number_Forms}     \p{Block=Number_Forms} (64)
   \p{Numeric_Type: De} \p{Numeric_Type=Decimal} (650)
   \p{Numeric_Type: Decimal} (Short: \p{Nt=De}) (650: [0-9],
                        U+0660..0669, U+06F0..06F9,
                        U+07C0..07C9, U+0966..096F, U+09E6..09EF
                        ...)
   \p{Numeric_Type: Di} \p{Numeric_Type=Digit} (128)
   \p{Numeric_Type: Digit} (Short: \p{Nt=Di}) (128: [\xb2-\xb3\xb9],
                        U+1369..1371, U+19DA, U+2070,
                        U+2074..2079, U+2080..2089 ...)
   \p{Numeric_Type: None}  (Short: \p{Nt=None}) (1_112_250 plus all
```

```
                    above-Unicode code points: [\x00-\x20!
                    \"#\$\%&\'\(\)*+,\-.\/:;<=>?\@A-Z\[\\\]
                    \^_'a-z\{\|\}~\x7f-\xb1\xb4-\xb8\xba-
                    \xbb\xbf-\xff], U+0100..065F,
                    U+066A..06EF, U+06FA..07BF,
                    U+07CA..0965, U+0970..09E5 ...)
```

  \p{Numeric_Type: Nu}   \p{Numeric_Type=Numeric} (1084)
  \p{Numeric_Type: Numeric} (Short: \p{Nt=Nu}) (1084: [\xbc-\xbe],
                    U+09F4..09F9, U+0B72..0B77,
                    U+0BF0..0BF2, U+0C78..0C7E, U+0D58..0D5E
                    ...)
T \p{Numeric_Value: -1/2} (Short: \p{Nv=-1/2}) (1: U+0F33)
T \p{Numeric_Value: 0}   (Short: \p{Nv=0}) (83: [0], U+0660,
                    U+06F0, U+07C0, U+0966, U+09E6 ...)
T \p{Numeric_Value: 1/320} (Short: \p{Nv=1/320}) (2: U+11FC0,
                    U+11FD4)
T \p{Numeric_Value: 1/160} (Short: \p{Nv=1/160}) (2: U+0D58, U+11FC1)
T \p{Numeric_Value: 1/80} (Short: \p{Nv=1/80}) (1: U+11FC2)
T \p{Numeric_Value: 1/64} (Short: \p{Nv=1/64}) (1: U+11FC3)
T \p{Numeric_Value: 1/40} (Short: \p{Nv=1/40}) (2: U+0D59, U+11FC4)
T \p{Numeric_Value: 1/32} (Short: \p{Nv=1/32}) (1: U+11FC5)
T \p{Numeric_Value: 3/80} (Short: \p{Nv=3/80}) (2: U+0D5A, U+11FC6)
T \p{Numeric_Value: 3/64} (Short: \p{Nv=3/64}) (1: U+11FC7)
T \p{Numeric_Value: 1/20} (Short: \p{Nv=1/20}) (2: U+0D5B, U+11FC8)
T \p{Numeric_Value: 1/16} (Short: \p{Nv=1/16}) (6: U+09F4, U+0B75,
                    U+0D76, U+A833, U+11FC9..11FCA)
T \p{Numeric_Value: 1/12} (Short: \p{Nv=1/12}) (1: U+109F6)
T \p{Numeric_Value: 1/10} (Short: \p{Nv=1/10}) (3: U+0D5C, U+2152,
                    U+11FCB)
T \p{Numeric_Value: 1/9} (Short: \p{Nv=1/9}) (1: U+2151)
T \p{Numeric_Value: 1/8} (Short: \p{Nv=1/8}) (7: U+09F5, U+0B76,
                    U+0D77, U+215B, U+A834, U+11FCC ...)
T \p{Numeric_Value: 1/7} (Short: \p{Nv=1/7}) (1: U+2150)
T \p{Numeric_Value: 3/20} (Short: \p{Nv=3/20}) (2: U+0D5D, U+11FCD)
T \p{Numeric_Value: 1/6} (Short: \p{Nv=1/6}) (4: U+2159, U+109F7,
                    U+12461, U+1ED3D)
T \p{Numeric_Value: 3/16} (Short: \p{Nv=3/16}) (5: U+09F6, U+0B77,
                    U+0D78, U+A835, U+11FCE)
T \p{Numeric_Value: 1/5} (Short: \p{Nv=1/5}) (3: U+0D5E, U+2155,
                    U+11FCF)

T \p{Numeric_Value: 1/4} (Short: \p{Nv=1/4}) (14: [\xbc], U+09F7,
              U+0B72, U+0D73, U+A830, U+10140 ...)
T \p{Numeric_Value: 1/3} (Short: \p{Nv=1/3}) (6: U+2153, U+109F9,
              U+10E7D, U+1245A, U+1245D, U+12465)
T \p{Numeric_Value: 3/8} (Short: \p{Nv=3/8}) (1: U+215C)
T \p{Numeric_Value: 2/5} (Short: \p{Nv=2/5}) (1: U+2156)
T \p{Numeric_Value: 5/12} (Short: \p{Nv=5/12}) (1: U+109FA)
T \p{Numeric_Value: 1/2} (Short: \p{Nv=1/2}) (19: [\xbd], U+0B73,
              U+0D74, U+0F2A, U+2CFD, U+A831 ...)
T \p{Numeric_Value: 7/12} (Short: \p{Nv=7/12}) (1: U+109FC)
T \p{Numeric_Value: 3/5} (Short: \p{Nv=3/5}) (1: U+2157)
T \p{Numeric_Value: 5/8} (Short: \p{Nv=5/8}) (1: U+215D)
T \p{Numeric_Value: 2/3} (Short: \p{Nv=2/3}) (7: U+2154, U+10177,
              U+109FD, U+10E7E, U+1245B, U+1245E ...)
T \p{Numeric_Value: 3/4} (Short: \p{Nv=3/4}) (9: [\xbe], U+09F8,
              U+0B74, U+0D75, U+A832, U+10178 ...)
T \p{Numeric_Value: 4/5} (Short: \p{Nv=4/5}) (1: U+2158)
T \p{Numeric_Value: 5/6} (Short: \p{Nv=5/6}) (3: U+215A, U+109FF,
              U+1245C)
T \p{Numeric_Value: 7/8} (Short: \p{Nv=7/8}) (1: U+215E)
T \p{Numeric_Value: 11/12} (Short: \p{Nv=11/12}) (1: U+109BC)
T \p{Numeric_Value: 1}   (Short: \p{Nv=1}) (140: [1\xb9], U+0661,
              U+06F1, U+07C1, U+0967, U+09E7 ...)
T \p{Numeric_Value: 3/2} (Short: \p{Nv=3/2}) (1: U+0F2B)
T \p{Numeric_Value: 2}   (Short: \p{Nv=2}) (139: [2\xb2], U+0662,
              U+06F2, U+07C2, U+0968, U+09E8 ...)
T \p{Numeric_Value: 5/2} (Short: \p{Nv=5/2}) (1: U+0F2C)
T \p{Numeric_Value: 3}   (Short: \p{Nv=3}) (140: [3\xb3], U+0663,
              U+06F3, U+07C3, U+0969, U+09E9 ...)
T \p{Numeric_Value: 7/2} (Short: \p{Nv=7/2}) (1: U+0F2D)
T \p{Numeric_Value: 4}   (Short: \p{Nv=4}) (131: [4], U+0664,
              U+06F4, U+07C4, U+096A, U+09EA ...)
T \p{Numeric_Value: 9/2} (Short: \p{Nv=9/2}) (1: U+0F2E)
T \p{Numeric_Value: 5}   (Short: \p{Nv=5}) (129: [5], U+0665,
              U+06F5, U+07C5, U+096B, U+09EB ...)
T \p{Numeric_Value: 11/2} (Short: \p{Nv=11/2}) (1: U+0F2F)
T \p{Numeric_Value: 6}   (Short: \p{Nv=6}) (113: [6], U+0666,
              U+06F6, U+07C6, U+096C, U+09EC ...)
T \p{Numeric_Value: 13/2} (Short: \p{Nv=13/2}) (1: U+0F30)
T \p{Numeric_Value: 7}   (Short: \p{Nv=7}) (112: [7], U+0667,

```
                    U+06F7, U+07C7, U+096D, U+09ED ...)
T \p{Numeric_Value: 15/2} (Short: \p{Nv=15/2}) (1: U+0F31)
T \p{Numeric_Value: 8}    (Short: \p{Nv=8}) (108: [8], U+0668,
                    U+06F8, U+07C8, U+096E, U+09EE ...)
T \p{Numeric_Value: 17/2} (Short: \p{Nv=17/2}) (1: U+0F32)
T \p{Numeric_Value: 9}    (Short: \p{Nv=9}) (112: [9], U+0669,
                    U+06F9, U+07C9, U+096F, U+09EF ...)
T \p{Numeric_Value: 10}   (Short: \p{Nv=10}) (62: U+0BF0, U+0D70,
                    U+1372, U+2169, U+2179, U+2469 ...)
T \p{Numeric_Value: 11}   (Short: \p{Nv=11}) (8: U+216A, U+217A,
                    U+246A, U+247E, U+2492, U+24EB ...)
T \p{Numeric_Value: 12}   (Short: \p{Nv=12}) (8: U+216B, U+217B,
                    U+246B, U+247F, U+2493, U+24EC ...)
T \p{Numeric_Value: 13}   (Short: \p{Nv=13}) (6: U+246C, U+2480,
                    U+2494, U+24ED, U+16E8D, U+1D2ED)
T \p{Numeric_Value: 14}   (Short: \p{Nv=14}) (6: U+246D, U+2481,
                    U+2495, U+24EE, U+16E8E, U+1D2EE)
T \p{Numeric_Value: 15}   (Short: \p{Nv=15}) (6: U+246E, U+2482,
                    U+2496, U+24EF, U+16E8F, U+1D2EF)
T \p{Numeric_Value: 16}   (Short: \p{Nv=16}) (7: U+09F9, U+246F,
                    U+2483, U+2497, U+24F0, U+16E90 ...)
T \p{Numeric_Value: 17}   (Short: \p{Nv=17}) (7: U+16EE, U+2470,
                    U+2484, U+2498, U+24F1, U+16E91 ...)
T \p{Numeric_Value: 18}   (Short: \p{Nv=18}) (7: U+16EF, U+2471,
                    U+2485, U+2499, U+24F2, U+16E92 ...)
T \p{Numeric_Value: 19}   (Short: \p{Nv=19}) (7: U+16F0, U+2472,
                    U+2486, U+249A, U+24F3, U+16E93 ...)
T \p{Numeric_Value: 20}   (Short: \p{Nv=20}) (36: U+1373, U+2473,
                    U+2487, U+249B, U+24F4, U+3039 ...)
T \p{Numeric_Value: 21}   (Short: \p{Nv=21}) (1: U+3251)
T \p{Numeric_Value: 22}   (Short: \p{Nv=22}) (1: U+3252)
T \p{Numeric_Value: 23}   (Short: \p{Nv=23}) (1: U+3253)
T \p{Numeric_Value: 24}   (Short: \p{Nv=24}) (1: U+3254)
T \p{Numeric_Value: 25}   (Short: \p{Nv=25}) (1: U+3255)
T \p{Numeric_Value: 26}   (Short: \p{Nv=26}) (1: U+3256)
T \p{Numeric_Value: 27}   (Short: \p{Nv=27}) (1: U+3257)
T \p{Numeric_Value: 28}   (Short: \p{Nv=28}) (1: U+3258)
T \p{Numeric_Value: 29}   (Short: \p{Nv=29}) (1: U+3259)
T \p{Numeric_Value: 30}   (Short: \p{Nv=30}) (19: U+1374, U+303A,
                    U+324A, U+325A, U+5345, U+10112 ...)
```

T \p{Numeric_Value: 31}   (Short: \p{Nv=31}) (1: U+325B)
T \p{Numeric_Value: 32}   (Short: \p{Nv=32}) (1: U+325C)
T \p{Numeric_Value: 33}   (Short: \p{Nv=33}) (1: U+325D)
T \p{Numeric_Value: 34}   (Short: \p{Nv=34}) (1: U+325E)
T \p{Numeric_Value: 35}   (Short: \p{Nv=35}) (1: U+325F)
T \p{Numeric_Value: 36}   (Short: \p{Nv=36}) (1: U+32B1)
T \p{Numeric_Value: 37}   (Short: \p{Nv=37}) (1: U+32B2)
T \p{Numeric_Value: 38}   (Short: \p{Nv=38}) (1: U+32B3)
T \p{Numeric_Value: 39}   (Short: \p{Nv=39}) (1: U+32B4)
T \p{Numeric_Value: 40}   (Short: \p{Nv=40}) (18: U+1375, U+324B,
                U+32B5, U+534C, U+10113, U+102ED ...)
T \p{Numeric_Value: 41}   (Short: \p{Nv=41}) (1: U+32B6)
T \p{Numeric_Value: 42}   (Short: \p{Nv=42}) (1: U+32B7)
T \p{Numeric_Value: 43}   (Short: \p{Nv=43}) (1: U+32B8)
T \p{Numeric_Value: 44}   (Short: \p{Nv=44}) (1: U+32B9)
T \p{Numeric_Value: 45}   (Short: \p{Nv=45}) (1: U+32BA)
T \p{Numeric_Value: 46}   (Short: \p{Nv=46}) (1: U+32BB)
T \p{Numeric_Value: 47}   (Short: \p{Nv=47}) (1: U+32BC)
T \p{Numeric_Value: 48}   (Short: \p{Nv=48}) (1: U+32BD)
T \p{Numeric_Value: 49}   (Short: \p{Nv=49}) (1: U+32BE)
T \p{Numeric_Value: 50}   (Short: \p{Nv=50}) (29: U+1376, U+216C,
                U+217C, U+2186, U+324C, U+32BF ...)
T \p{Numeric_Value: 60}   (Short: \p{Nv=60}) (13: U+1377, U+324D,
                U+10115, U+102EF, U+109CE, U+10E6E ...)
T \p{Numeric_Value: 70}   (Short: \p{Nv=70}) (13: U+1378, U+324E,
                U+10116, U+102F0, U+109CF, U+10E6F ...)
T \p{Numeric_Value: 80}   (Short: \p{Nv=80}) (12: U+1379, U+324F,
                U+10117, U+102F1, U+10E70, U+11062 ...)
T \p{Numeric_Value: 90}   (Short: \p{Nv=90}) (12: U+137A, U+10118,
                U+102F2, U+10341, U+10E71, U+11063 ...)
T \p{Numeric_Value: 100} (Short: \p{Nv=100}) (35: U+0BF1, U+0D71,
                U+137B, U+216D, U+217D, U+4F70 ...)
T \p{Numeric_Value: 200} (Short: \p{Nv=200}) (6: U+1011A, U+102F4,
                U+109D3, U+10E73, U+1EC84, U+1ED14)
T \p{Numeric_Value: 300} (Short: \p{Nv=300}) (7: U+1011B, U+1016B,
                U+102F5, U+109D4, U+10E74, U+1EC85 ...)
T \p{Numeric_Value: 400} (Short: \p{Nv=400}) (7: U+1011C, U+102F6,
                U+109D5, U+10E75, U+1EC86, U+1ED16 ...)
T \p{Numeric_Value: 500} (Short: \p{Nv=500}) (16: U+216E, U+217E,
                U+1011D, U+10145, U+1014C, U+10153 ...)

T \p{Numeric_Value: 600} (Short: \p{Nv=600}) (7: U+1011E, U+102F8,
            U+109D7, U+10E77, U+1EC88, U+1ED18 ...)
T \p{Numeric_Value: 700} (Short: \p{Nv=700}) (6: U+1011F, U+102F9,
            U+109D8, U+10E78, U+1EC89, U+1ED19)
T \p{Numeric_Value: 800} (Short: \p{Nv=800}) (6: U+10120, U+102FA,
            U+109D9, U+10E79, U+1EC8A, U+1ED1A)
T \p{Numeric_Value: 900} (Short: \p{Nv=900}) (7: U+10121, U+102FB,
            U+1034A, U+109DA, U+10E7A, U+1EC8B ...)
T \p{Numeric_Value: 1000} (Short: \p{Nv=1000}) (22: U+0BF2, U+0D72,
            U+216F, U+217F..2180, U+4EDF, U+5343 ...)
T \p{Numeric_Value: 2000} (Short: \p{Nv=2000}) (5: U+10123, U+109DC,
            U+1EC8D, U+1ED1D, U+1ED3A)
T \p{Numeric_Value: 3000} (Short: \p{Nv=3000}) (4: U+10124, U+109DD,
            U+1EC8E, U+1ED1E)
T \p{Numeric_Value: 4000} (Short: \p{Nv=4000}) (4: U+10125, U+109DE,
            U+1EC8F, U+1ED1F)
T \p{Numeric_Value: 5000} (Short: \p{Nv=5000}) (8: U+2181, U+10126,
            U+10146, U+1014E, U+10172, U+109DF ...)
T \p{Numeric_Value: 6000} (Short: \p{Nv=6000}) (4: U+10127, U+109E0,
            U+1EC91, U+1ED21)
T \p{Numeric_Value: 7000} (Short: \p{Nv=7000}) (4: U+10128, U+109E1,
            U+1EC92, U+1ED22)
T \p{Numeric_Value: 8000} (Short: \p{Nv=8000}) (4: U+10129, U+109E2,
            U+1EC93, U+1ED23)
T \p{Numeric_Value: 9000} (Short: \p{Nv=9000}) (4: U+1012A, U+109E3,
            U+1EC94, U+1ED24)
T \p{Numeric_Value: 10000} (= 1.0e+04) (Short: \p{Nv=10000}) (13:
            U+137C, U+2182, U+4E07, U+842C, U+1012B,
            U+10155 ...)
T \p{Numeric_Value: 20000} (= 2.0e+04) (Short: \p{Nv=20000}) (4:
            U+1012C, U+109E5, U+1EC96, U+1ED26)
T \p{Numeric_Value: 30000} (= 3.0e+04) (Short: \p{Nv=30000}) (4:
            U+1012D, U+109E6, U+1EC97, U+1ED27)
T \p{Numeric_Value: 40000} (= 4.0e+04) (Short: \p{Nv=40000}) (4:
            U+1012E, U+109E7, U+1EC98, U+1ED28)
T \p{Numeric_Value: 50000} (= 5.0e+04) (Short: \p{Nv=50000}) (7:
            U+2187, U+1012F, U+10147, U+10156,
            U+109E8, U+1EC99 ...)
T \p{Numeric_Value: 60000} (= 6.0e+04) (Short: \p{Nv=60000}) (4:
            U+10130, U+109E9, U+1EC9A, U+1ED2A)

T \p{Numeric_Value: 70000} (= 7.0e+04) (Short: \p{Nv=70000}) (4:
            U+10131, U+109EA, U+1EC9B, U+1ED2B)

T \p{Numeric_Value: 80000} (= 8.0e+04) (Short: \p{Nv=80000}) (4:
            U+10132, U+109EB, U+1EC9C, U+1ED2C)

T \p{Numeric_Value: 90000} (= 9.0e+04) (Short: \p{Nv=90000}) (4:
            U+10133, U+109EC, U+1EC9D, U+1ED2D)

T \p{Numeric_Value: 100000} (= 1.0e+05) (Short: \p{Nv=100000}) (5:
            U+2188, U+109ED, U+1EC9E, U+1ECA0,
            U+1ECB4)

T \p{Numeric_Value: 200000} (= 2.0e+05) (Short: \p{Nv=200000}) (2:
            U+109EE, U+1EC9F)

T \p{Numeric_Value: 216000} (= 2.2e+05) (Short: \p{Nv=216000}) (1:
            U+12432)

T \p{Numeric_Value: 300000} (= 3.0e+05) (Short: \p{Nv=300000}) (1:
            U+109EF)

T \p{Numeric_Value: 400000} (= 4.0e+05) (Short: \p{Nv=400000}) (1:
            U+109F0)

T \p{Numeric_Value: 432000} (= 4.3e+05) (Short: \p{Nv=432000}) (1:
            U+12433)

T \p{Numeric_Value: 500000} (= 5.0e+05) (Short: \p{Nv=500000}) (1:
            U+109F1)

T \p{Numeric_Value: 600000} (= 6.0e+05) (Short: \p{Nv=600000}) (1:
            U+109F2)

T \p{Numeric_Value: 700000} (= 7.0e+05) (Short: \p{Nv=700000}) (1:
            U+109F3)

T \p{Numeric_Value: 800000} (= 8.0e+05) (Short: \p{Nv=800000}) (1:
            U+109F4)

T \p{Numeric_Value: 900000} (= 9.0e+05) (Short: \p{Nv=900000}) (1:
            U+109F5)

T \p{Numeric_Value: 1000000} (= 1.0e+06) (Short: \p{Nv=1000000}) (1:
            U+16B5E)

T \p{Numeric_Value: 10000000} (= 1.0e+07) (Short: \p{Nv=10000000})
            (1: U+1ECA1)

T \p{Numeric_Value: 20000000} (= 2.0e+07) (Short: \p{Nv=20000000})
            (1: U+1ECA2)

T \p{Numeric_Value: 100000000} (= 1.0e+08) (Short: \p{Nv=100000000})
            (3: U+4EBF, U+5104, U+16B5F)

T \p{Numeric_Value: 10000000000} (= 1.0e+10) (Short: \p{Nv=
            10000000000}) (1: U+16B60)

T \p{Numeric_Value: 1000000000000} (= 1.0e+12) (Short: \p{Nv=

                      1000000000000}) (2: U+5146, U+16B61)
   \p{Numeric_Value: NaN} (Short: \p{Nv=NaN}) (1_112_250 plus all
                      above-Unicode code points: [\x00-\x20!
                      \"#\$\%&\'\(\)*+,\-.\/:;<=>?\@A-Z\[\\\]
                      \^_'a-z\{\|\}~\x7f-\xb1\xb4-\xb8\xba-
                      \xbb\xbf-\xff], U+0100..065F,
                      U+066A..06EF, U+06FA..07BF,
                      U+07CA..0965, U+0970..09E5 ...)
   \p{Nushu}   \p{Script_Extensions=Nushu} (Short:
                      \p{Nshu}; NOT \p{Block=Nushu}) (397)
   \p{Nv: *}   \p{Numeric_Value: *}
   \p{Nyiakeng_Puachue_Hmong} \p{Script_Extensions=
                      Nyiakeng_Puachue_Hmong} (Short:
                      \p{Hmnp}; NOT \p{Block=
                      Nyiakeng_Puachue_Hmong}) (71)
 X \p{OCR}   \p{Optical_Character_Recognition} (=
                      \p{Block=Optical_Character_Recognition})
                      (32)
   \p{Ogam}   \p{Ogham} (= \p{Script_Extensions=Ogham})
                      (NOT \p{Block=Ogham}) (29)
   \p{Ogham}   \p{Script_Extensions=Ogham} (Short:
                      \p{Ogam}; NOT \p{Block=Ogham}) (29)
   \p{Ol_Chiki}   \p{Script_Extensions=Ol_Chiki} (Short:
                      \p{Olck}) (48)
   \p{Olck}   \p{Ol_Chiki} (= \p{Script_Extensions=
                      Ol_Chiki}) (48)
   \p{Old_Hungarian} \p{Script_Extensions=Old_Hungarian}
                      (Short: \p{Hung}; NOT \p{Block=
                      Old_Hungarian}) (108)
   \p{Old_Italic}   \p{Script_Extensions=Old_Italic} (Short:
                      \p{Ital}; NOT \p{Block=Old_Italic}) (39)
   \p{Old_North_Arabian} \p{Script_Extensions=Old_North_Arabian}
                      (Short: \p{Narb}) (32)
   \p{Old_Permic}   \p{Script_Extensions=Old_Permic} (Short:
                      \p{Perm}; NOT \p{Block=Old_Permic}) (44)
   \p{Old_Persian}   \p{Script_Extensions=Old_Persian} (Short:
                      \p{Xpeo}; NOT \p{Block=Old_Persian}) (50)
   \p{Old_Sogdian}   \p{Script_Extensions=Old_Sogdian} (Short:
                      \p{Sogo}; NOT \p{Block=Old_Sogdian}) (40)
   \p{Old_South_Arabian} \p{Script_Extensions=Old_South_Arabian}

(Short: \p{Sarb}) (32)

\p{Old_Turkic}   \p{Script_Extensions=Old_Turkic} (Short:
                \p{Orkh}; NOT \p{Block=Old_Turkic}) (73)

\p{Open_Punctuation}   \p{General_Category=Open_Punctuation}
                (Short: \p{Ps}) (75)

X \p{Optical_Character_Recognition} \p{Block=
                Optical_Character_Recognition} (Short:
                \p{InOCR}) (32)

\p{Oriya}   \p{Script_Extensions=Oriya} (Short:
                \p{Orya}; NOT \p{Block=Oriya}) (97)

\p{Orkh}   \p{Old_Turkic} (= \p{Script_Extensions=
                Old_Turkic}) (NOT \p{Block=Old_Turkic})
                (73)

X \p{Ornamental_Dingbats} \p{Block=Ornamental_Dingbats} (48)

\p{Orya}   \p{Oriya} (= \p{Script_Extensions=Oriya})
                (NOT \p{Block=Oriya}) (97)

\p{Osage}   \p{Script_Extensions=Osage} (Short:
                \p{Osge}; NOT \p{Block=Osage}) (72)

\p{Osge}   \p{Osage} (= \p{Script_Extensions=Osage})
                (NOT \p{Block=Osage}) (72)

\p{Osma}   \p{Osmanya} (= \p{Script_Extensions=
                Osmanya}) (NOT \p{Block=Osmanya}) (40)

\p{Osmanya}   \p{Script_Extensions=Osmanya} (Short:
                \p{Osma}; NOT \p{Block=Osmanya}) (40)

\p{Other}   \p{General_Category=Other} (Short: \p{C})
                (970_414 plus all above-Unicode code
                points)

\p{Other_Letter}   \p{General_Category=Other_Letter} (Short:
                \p{Lo}) (127_004)

\p{Other_Number}   \p{General_Category=Other_Number} (Short:
                \p{No}) (895)

\p{Other_Punctuation}   \p{General_Category=Other_Punctuation}
                (Short: \p{Po}) (593)

\p{Other_Symbol}   \p{General_Category=Other_Symbol} (Short:
                \p{So}) (6431)

X \p{Ottoman_Siyaq_Numbers} \p{Block=Ottoman_Siyaq_Numbers} (80)

\p{P} \pP   \p{Punct} (= \p{General_Category=
                Punctuation}) (NOT
                \p{General_Punctuation}) (798)

\p{Pahawh_Hmong}   \p{Script_Extensions=Pahawh_Hmong} (Short:

```
                    \p{Hmng}; NOT \p{Block=Pahawh_Hmong})
                    (127)
\p{Palm}            \p{Palmyrene} (= \p{Script_Extensions=
                    Palmyrene}) (32)
\p{Palmyrene}        \p{Script_Extensions=Palmyrene} (Short:
                    \p{Palm}) (32)
\p{Paragraph_Separator} \p{General_Category=Paragraph_Separator}
                    (Short: \p{Zp}) (1)
\p{Pat_Syn}          \p{Pattern_Syntax} (= \p{Pattern_Syntax=
                    Y}) (2760)
\p{Pat_Syn: *}        \p{Pattern_Syntax: *}
\p{Pat_WS}           \p{Pattern_White_Space} (=
                    \p{Pattern_White_Space=Y}) (11)
\p{Pat_WS: *}         \p{Pattern_White_Space: *}
\p{Pattern_Syntax}    \p{Pattern_Syntax=Y} (Short: \p{PatSyn})
                    (2760)
\p{Pattern_Syntax: N*}  (Short: \p{PatSyn=N}, \P{PatSyn})
                    (1_111_352 plus all above-Unicode code
                    points: [\x00-\x200-9A-Z_a-z\x7f-\xa0
                    \xa8\xaa\xad\xaf\xb2-\xb5\xb7-\xba\xbc-
                    \xbe\xc0-\xd6\xd8-\xf6\xf8-\xff],
                    U+0100..200F, U+2028..202F,
                    U+203F..2040, U+2054, U+205F..218F ...)
\p{Pattern_Syntax: Y*}  (Short: \p{PatSyn=Y}, \p{PatSyn}) (2760:
                    [!\"#\$\%&\'\(\)*+,\-.\/:;<=>?\@\[\\\]
                    \^`\{\|\}~\xa1-\xa7\xa9\xab-\xac\xae
                    \xb0-\xb1\xb6\xbb\xbf\xd7\xf7],
                    U+2010..2027, U+2030..203E,
                    U+2041..2053, U+2055..205E, U+2190..245F
                    ...)
\p{Pattern_White_Space} \p{Pattern_White_Space=Y} (Short:
                    \p{PatWS}) (11)
\p{Pattern_White_Space: N*} (Short: \p{PatWS=N}, \P{PatWS})
                    (1_114_101 plus all above-Unicode code
                    points: [^\t\n\cK\f\r\x20\x85],
                    U+0100..200D, U+2010..2027,
                    U+202A..infinity)
\p{Pattern_White_Space: Y*} (Short: \p{PatWS=Y}, \p{PatWS}) (11:
                    [\t\n\cK\f\r\x20\x85], U+200E..200F,
                    U+2028..2029)
```

\p{Pau_Cin_Hau}   \p{Script_Extensions=Pau_Cin_Hau} (Short:
              \p{Pauc}; NOT \p{Block=Pau_Cin_Hau}) (57)

\p{Pauc}   \p{Pau_Cin_Hau} (= \p{Script_Extensions=
              Pau_Cin_Hau}) (NOT \p{Block=
              Pau_Cin_Hau}) (57)

\p{Pc}   \p{Connector_Punctuation} (=
              \p{General_Category=
              Connector_Punctuation}) (10)

\p{PCM}   \p{Prepended_Concatenation_Mark} (=
              \p{Prepended_Concatenation_Mark=Y}) (11)

\p{PCM: *}   \p{Prepended_Concatenation_Mark: *}

\p{Pd}   \p{Dash_Punctuation} (=
              \p{General_Category=Dash_Punctuation})
              (25)

\p{Pe}   \p{Close_Punctuation} (=
              \p{General_Category=Close_Punctuation})
              (73)

\p{PerlSpace}   \p{PosixSpace} (6)

\p{PerlWord}   \p{PosixWord} (63)

\p{Perm}   \p{Old_Permic} (= \p{Script_Extensions=
              Old_Permic}) (NOT \p{Block=Old_Permic})
              (44)

\p{Pf}   \p{Final_Punctuation} (=
              \p{General_Category=Final_Punctuation})
              (10)

\p{Phag}   \p{Phags_Pa} (= \p{Script_Extensions=
              Phags_Pa}) (NOT \p{Block=Phags_Pa}) (59)

\p{Phags_Pa}   \p{Script_Extensions=Phags_Pa} (Short:
              \p{Phag}; NOT \p{Block=Phags_Pa}) (59)

X \p{Phaistos}   \p{Phaistos_Disc} (= \p{Block=
              Phaistos_Disc}) (48)

X \p{Phaistos_Disc}   \p{Block=Phaistos_Disc} (Short:
              \p{InPhaistos}) (48)

\p{Phli}   \p{Inscriptional_Pahlavi} (=
              \p{Script_Extensions=
              Inscriptional_Pahlavi}) (NOT \p{Block=
              Inscriptional_Pahlavi}) (27)

\p{Phlp}   \p{Psalter_Pahlavi} (=
              \p{Script_Extensions=Psalter_Pahlavi})
              (NOT \p{Block=Psalter_Pahlavi}) (30)

```
    \p{Phnx}              \p{Phoenician} (= \p{Script_Extensions=
                          Phoenician}) (NOT \p{Block=Phoenician})
                          (29)
    \p{Phoenician}        \p{Script_Extensions=Phoenician} (Short:
                          \p{Phnx}; NOT \p{Block=Phoenician}) (29)
X \p{Phonetic_Ext}       \p{Phonetic_Extensions} (= \p{Block=
                          Phonetic_Extensions}) (128)
X \p{Phonetic_Ext_Sup}   \p{Phonetic_Extensions_Supplement} (=
                          \p{Block=
                          Phonetic_Extensions_Supplement}) (64)
X \p{Phonetic_Extensions} \p{Block=Phonetic_Extensions} (Short:
                          \p{InPhoneticExt}) (128)
X \p{Phonetic_Extensions_Supplement} \p{Block=
                          Phonetic_Extensions_Supplement} (Short:
                          \p{InPhoneticExtSup}) (64)
    \p{Pi}                \p{Initial_Punctuation} (=
                          \p{General_Category=
                          Initial_Punctuation}) (12)
X \p{Playing_Cards}      \p{Block=Playing_Cards} (96)
    \p{Plrd}              \p{Miao} (= \p{Script_Extensions=Miao})
                          (NOT \p{Block=Miao}) (149)
    \p{Po}                \p{Other_Punctuation} (=
                          \p{General_Category=Other_Punctuation})
                          (593)
    \p{PosixAlnum}        (62: [0-9A-Za-z])
    \p{PosixAlpha}        (52: [A-Za-z])
    \p{PosixBlank}        (2: [\t\x20])
    \p{PosixCntrl}        ASCII control characters (33: ACK, BEL,
                          BS, CAN, CR, DC1, DC2, DC3, DC4, DEL,
                          DLE, ENQ, EOM, EOT, ESC, ETB, ETX, FF,
                          FS, GS, HT, LF, NAK, NUL, RS, SI, SO,
                          SOH, STX, SUB, SYN, US, VT)
    \p{PosixDigit}        (10: [0-9])
    \p{PosixGraph}        (94: [!\"#\$\%&\'\(\)*+,\-.\/0-9:;<=>?\@A-
                          Z\[\\\]\^_`a-z\{\|\}~])
    \p{PosixLower}        (/i= PosixAlpha) (26: [a-z])
    \p{PosixPrint}        (95: [\x20-\x7e])
    \p{PosixPunct}        (32: [!\"#\$\%&\'\(\)*+,\-.\/:;<=>?\@
                          \[\\\]\^_`\{\|\}~])
    \p{PosixSpace}        (Short: \p{PerlSpace}) (6: [\t\n\cK\f\r
```

```
                        \x20])
    \p{PosixUpper}       (/i= PosixAlpha) (26: [A-Z])
    \p{PosixWord}        \w, restricted to ASCII (Short:
                        \p{PerlWord}) (63: [0-9A-Z_a-z])
    \p{PosixXDigit}      \p{ASCII_Hex_Digit=Y} (Short: \p{AHex})
                        (22)
    \p{Prepended_Concatenation_Mark} \p{Prepended_Concatenation_Mark=
                        Y} (Short: \p{PCM}) (11)
    \p{Prepended_Concatenation_Mark: N*} (Short: \p{PCM=N}, \P{PCM})
                        (1_114_101 plus all above-Unicode code
                        points: U+0000..05FF, U+0606..06DC,
                        U+06DE..070E, U+0710..08E1,
                        U+08E3..110BC, U+110BE..110CC ...)
    \p{Prepended_Concatenation_Mark: Y*} (Short: \p{PCM=Y}, \p{PCM})
                        (11: U+0600..0605, U+06DD, U+070F,
                        U+08E2, U+110BD, U+110CD)
  T \p{Present_In: 1.1}    \p{Age=V1_1} (Short: \p{In=1.1}) (Perl
                        extension) (33_979)
  T \p{Present_In: 2.0}    Code point's usage introduced in version
                        2.0 or earlier (Short: \p{In=2.0}) (Perl
                        extension) (178_500: U+0000..01F5,
                        U+01FA..0217, U+0250..02A8,
                        U+02B0..02DE, U+02E0..02E9, U+0300..0345
                        ...)
    \p{Present_In: V2_0}    \p{Present_In=2.0} (Perl extension)
                        (178_500)
  T \p{Present_In: 2.1}    Code point's usage introduced in version
                        2.1 or earlier (Short: \p{In=2.1}) (Perl
                        extension) (178_502: U+0000..01F5,
                        U+01FA..0217, U+0250..02A8,
                        U+02B0..02DE, U+02E0..02E9, U+0300..0345
                        ...)
    \p{Present_In: V2_1}    \p{Present_In=2.1} (Perl extension)
                        (178_502)
  T \p{Present_In: 3.0}    Code point's usage introduced in version
                        3.0 or earlier (Short: \p{In=3.0}) (Perl
                        extension) (188_809: U+0000..021F,
                        U+0222..0233, U+0250..02AD,
                        U+02B0..02EE, U+0300..034E, U+0360..0362
                        ...)
```

```
   \p{Present_In: V3_0}   \p{Present_In=3.0} (Perl extension)
                         (188_809)
T \p{Present_In: 3.1}    Code point's usage introduced in version
                         3.1 or earlier (Short: \p{In=3.1}) (Perl
                         extension) (233_787: U+0000..021F,
                         U+0222..0233, U+0250..02AD,
                         U+02B0..02EE, U+0300..034E, U+0360..0362
                         ...)
   \p{Present_In: V3_1}   \p{Present_In=3.1} (Perl extension)
                         (233_787)
T \p{Present_In: 3.2}    Code point's usage introduced in version
                         3.2 or earlier (Short: \p{In=3.2}) (Perl
                         extension) (234_803: U+0000..0220,
                         U+0222..0233, U+0250..02AD,
                         U+02B0..02EE, U+0300..034F, U+0360..036F
                         ...)
   \p{Present_In: V3_2}   \p{Present_In=3.2} (Perl extension)
                         (234_803)
T \p{Present_In: 4.0}    Code point's usage introduced in version
                         4.0 or earlier (Short: \p{In=4.0}) (Perl
                         extension) (236_029: U+0000..0236,
                         U+0250..0357, U+035D..036F,
                         U+0374..0375, U+037A, U+037E ...)
   \p{Present_In: V4_0}   \p{Present_In=4.0} (Perl extension)
                         (236_029)
T \p{Present_In: 4.1}    Code point's usage introduced in version
                         4.1 or earlier (Short: \p{In=4.1}) (Perl
                         extension) (237_302: U+0000..0241,
                         U+0250..036F, U+0374..0375, U+037A,
                         U+037E, U+0384..038A ...)
   \p{Present_In: V4_1}   \p{Present_In=4.1} (Perl extension)
                         (237_302)
T \p{Present_In: 5.0}    Code point's usage introduced in version
                         5.0 or earlier (Short: \p{In=5.0}) (Perl
                         extension) (238_671: U+0000..036F,
                         U+0374..0375, U+037A..037E,
                         U+0384..038A, U+038C, U+038E..03A1 ...)
   \p{Present_In: V5_0}   \p{Present_In=5.0} (Perl extension)
                         (238_671)
T \p{Present_In: 5.1}    Code point's usage introduced in version
```

5.1 or earlier (Short: \p{In=5.1}) (Perl
extension) (240_295: U+0000..0377,
U+037A..037E, U+0384..038A, U+038C,
U+038E..03A1, U+03A3..0523 ...)

\p{Present_In: V5_1}   \p{Present_In=5.1} (Perl extension)
(240_295)

T \p{Present_In: 5.2}   Code point's usage introduced in version
5.2 or earlier (Short: \p{In=5.2}) (Perl
extension) (246_943: U+0000..0377,
U+037A..037E, U+0384..038A, U+038C,
U+038E..03A1, U+03A3..0525 ...)

\p{Present_In: V5_2}   \p{Present_In=5.2} (Perl extension)
(246_943)

T \p{Present_In: 6.0}   Code point's usage introduced in version
6.0 or earlier (Short: \p{In=6.0}) (Perl
extension) (249_031: U+0000..0377,
U+037A..037E, U+0384..038A, U+038C,
U+038E..03A1, U+03A3..0527 ...)

\p{Present_In: V6_0}   \p{Present_In=6.0} (Perl extension)
(249_031)

T \p{Present_In: 6.1}   Code point's usage introduced in version
6.1 or earlier (Short: \p{In=6.1}) (Perl
extension) (249_763: U+0000..0377,
U+037A..037E, U+0384..038A, U+038C,
U+038E..03A1, U+03A3..0527 ...)

\p{Present_In: V6_1}   \p{Present_In=6.1} (Perl extension)
(249_763)

T \p{Present_In: 6.2}   Code point's usage introduced in version
6.2 or earlier (Short: \p{In=6.2}) (Perl
extension) (249_764: U+0000..0377,
U+037A..037E, U+0384..038A, U+038C,
U+038E..03A1, U+03A3..0527 ...)

\p{Present_In: V6_2}   \p{Present_In=6.2} (Perl extension)
(249_764)

T \p{Present_In: 6.3}   Code point's usage introduced in version
6.3 or earlier (Short: \p{In=6.3}) (Perl
extension) (249_769: U+0000..0377,
U+037A..037E, U+0384..038A, U+038C,
U+038E..03A1, U+03A3..0527 ...)

\p{Present_In: V6_3}   \p{Present_In=6.3} (Perl extension)

                 (249_769)
    T \p{Present_In: 7.0}   Code point's usage introduced in version
                 7.0 or earlier (Short: \p{In=7.0}) (Perl
                 extension) (252_603: U+0000..0377,
                 U+037A..037F, U+0384..038A, U+038C,
                 U+038E..03A1, U+03A3..052F ...)
     \p{Present_In: V7_0}   \p{Present_In=7.0} (Perl extension)
                 (252_603)
    T \p{Present_In: 8.0}   Code point's usage introduced in version
                 8.0 or earlier (Short: \p{In=8.0}) (Perl
                 extension) (260_319: U+0000..0377,
                 U+037A..037F, U+0384..038A, U+038C,
                 U+038E..03A1, U+03A3..052F ...)
     \p{Present_In: V8_0}   \p{Present_In=8.0} (Perl extension)
                 (260_319)
    T \p{Present_In: 9.0}   Code point's usage introduced in version
                 9.0 or earlier (Short: \p{In=9.0}) (Perl
                 extension) (267_819: U+0000..0377,
                 U+037A..037F, U+0384..038A, U+038C,
                 U+038E..03A1, U+03A3..052F ...)
     \p{Present_In: V9_0}   \p{Present_In=9.0} (Perl extension)
                 (267_819)
    T \p{Present_In: 10.0}   Code point's usage introduced in version
                 10.0 or earlier (Short: \p{In=10.0})
                 (Perl extension) (276_337: U+0000..0377,
                 U+037A..037F, U+0384..038A, U+038C,
                 U+038E..03A1, U+03A3..052F ...)
     \p{Present_In: V10_0} \p{Present_In=10.0} (Perl extension)
                 (276_337)
    T \p{Present_In: 11.0}   Code point's usage introduced in version
                 11.0 or earlier (Short: \p{In=11.0})
                 (Perl extension) (277_021: U+0000..0377,
                 U+037A..037F, U+0384..038A, U+038C,
                 U+038E..03A1, U+03A3..052F ...)
     \p{Present_In: V11_0} \p{Present_In=11.0} (Perl extension)
                 (277_021)
    T \p{Present_In: 12.0}   Code point's usage introduced in version
                 12.0 or earlier (Short: \p{In=12.0})
                 (Perl extension) (277_575: U+0000..0377,
                 U+037A..037F, U+0384..038A, U+038C,

                      U+038E..03A1, U+03A3..052F ...)
    \p{Present_In: V12_0} \p{Present_In=12.0} (Perl extension)
                      (277_575)
  T \p{Present_In: 12.1}   Code point's usage introduced in version
                      12.1 or earlier (Short: \p{In=12.1})
                      (Perl extension) (277_576: U+0000..0377,
                      U+037A..037F, U+0384..038A, U+038C,
                      U+038E..03A1, U+03A3..052F ...)
    \p{Present_In: V12_1} \p{Present_In=12.1} (Perl extension)
                      (277_576)
  T \p{Present_In: 13.0}   Code point's usage introduced in version
                      13.0 or earlier (Short: \p{In=13.0})
                      (Perl extension) (283_506: U+0000..0377,
                      U+037A..037F, U+0384..038A, U+038C,
                      U+038E..03A1, U+03A3..052F ...)
    \p{Present_In: V13_0} \p{Present_In=13.0} (Perl extension)
                      (283_506)
    \p{Present_In: Unassigned} \p{Age=Unassigned} (Short: \p{In=
                      Unassigned}) (Perl extension) (830_606
                      plus all above-Unicode code points)
    \p{Print}   \p{XPosixPrint} (281_325)
    \p{Private_Use}   \p{General_Category=Private_Use} (Short:
                      \p{Co}; NOT \p{Private_Use_Area})
                      (137_468)
  X \p{Private_Use_Area}   \p{Block=Private_Use_Area} (Short:
                      \p{InPUA}) (6400)
    \p{Prti}   \p{Inscriptional_Parthian} (=
                      \p{Script_Extensions=
                      Inscriptional_Parthian}) (NOT \p{Block=
                      Inscriptional_Parthian}) (30)
    \p{Ps}   \p{Open_Punctuation} (=
                      \p{General_Category=Open_Punctuation})
                      (75)
    \p{Psalter_Pahlavi}   \p{Script_Extensions=Psalter_Pahlavi}
                      (Short: \p{Phlp}; NOT \p{Block=
                      Psalter_Pahlavi}) (30)
  X \p{PUA}   \p{Private_Use_Area} (= \p{Block=
                      Private_Use_Area}) (6400)
    \p{Punct}   \p{General_Category=Punctuation} (Short:
                      \p{P}; NOT \p{General_Punctuation}) (798)

\p{Punctuation}   \p{Punct} (= \p{General_Category=
                Punctuation}) (NOT
                \p{General_Punctuation}) (798)
\p{Qaac}   \p{Coptic} (= \p{Script_Extensions=
                Coptic}) (NOT \p{Block=Coptic}) (165)
\p{Qaai}   \p{Inherited} (= \p{Script_Extensions=
                Inherited}) (503)
\p{QMark}   \p{Quotation_Mark} (= \p{Quotation_Mark=
                Y}) (30)
\p{QMark: *}   \p{Quotation_Mark: *}
\p{Quotation_Mark}   \p{Quotation_Mark=Y} (Short: \p{QMark})
                (30)
\p{Quotation_Mark: N*} (Short: \p{QMark=N}, \P{QMark}) (1_114_082
                plus all above-Unicode code points:
                [\x00-\x20!#\$\%&\(\)*+,\-.\/0-9:;<=>?
                \@A-Z\[\\\]\^_'a-z\{\|\}~\x7f-\xaa\xac-
                \xba\xbc-\xff], U+0100..2017,
                U+2020..2038, U+203B..2E41,
                U+2E43..300B, U+3010..301C ...)
\p{Quotation_Mark: Y*} (Short: \p{QMark=Y}, \p{QMark}) (30: [\"
                \'\xab\xbb], U+2018..201F, U+2039..203A,
                U+2E42, U+300C..300F, U+301D..301F ...)
\p{Radical}   \p{Radical=Y} (329)
\p{Radical: N*}   (Single: \P{Radical}) (1_113_783 plus all
                above-Unicode code points: U+0000..2E7F,
                U+2E9A, U+2EF4..2EFF, U+2FD6..infinity)
\p{Radical: Y*}   (Single: \p{Radical}) (329: U+2E80..2E99,
                U+2E9B..2EF3, U+2F00..2FD5)
\p{Regional_Indicator} \p{Regional_Indicator=Y} (Short: \p{RI})
                (26)
\p{Regional_Indicator: N*} (Short: \p{RI=N}, \P{RI}) (1_114_086
                plus all above-Unicode code points:
                U+0000..1F1E5, U+1F200..infinity)
\p{Regional_Indicator: Y*} (Short: \p{RI=Y}, \p{RI}) (26:
                U+1F1E6..1F1FF)
\p{Rejang}   \p{Script_Extensions=Rejang} (Short:
                \p{Rjng}; NOT \p{Block=Rejang}) (37)
\p{RI}   \p{Regional_Indicator} (=
                \p{Regional_Indicator=Y}) (26)
\p{RI: *}   \p{Regional_Indicator: *}

```
 \p{Rjng}            \p{Rejang} (= \p{Script_Extensions=
                     Rejang}) (NOT \p{Block=Rejang}) (37)
 \p{Rohg}            \p{Hanifi_Rohingya} (=
                     \p{Script_Extensions=Hanifi_Rohingya})
                     (NOT \p{Block=Hanifi_Rohingya}) (55)
X \p{Rumi}           \p{Rumi_Numeral_Symbols} (= \p{Block=
                     Rumi_Numeral_Symbols}) (32)
X \p{Rumi_Numeral_Symbols} \p{Block=Rumi_Numeral_Symbols} (Short:
                     \p{InRumi}) (32)
 \p{Runic}           \p{Script_Extensions=Runic} (Short:
                     \p{Runr}; NOT \p{Block=Runic}) (86)
 \p{Runr}            \p{Runic} (= \p{Script_Extensions=Runic})
                     (NOT \p{Block=Runic}) (86)
 \p{S} \pS           \p{Symbol} (= \p{General_Category=Symbol})
                     (7564)
 \p{Samaritan}       \p{Script_Extensions=Samaritan} (Short:
                     \p{Samr}; NOT \p{Block=Samaritan}) (61)
 \p{Samr}            \p{Samaritan} (= \p{Script_Extensions=
                     Samaritan}) (NOT \p{Block=Samaritan})
                     (61)
 \p{Sarb}            \p{Old_South_Arabian} (=
                     \p{Script_Extensions=Old_South_Arabian})
                     (32)
 \p{Saur}            \p{Saurashtra} (= \p{Script_Extensions=
                     Saurashtra}) (NOT \p{Block=Saurashtra})
                     (82)
 \p{Saurashtra}      \p{Script_Extensions=Saurashtra} (Short:
                     \p{Saur}; NOT \p{Block=Saurashtra}) (82)
 \p{SB: *}           \p{Sentence_Break: *}
 \p{Sc}              \p{Currency_Symbol} (=
                     \p{General_Category=Currency_Symbol})
                     (62)
 \p{Sc: *}           \p{Script: *}
 \p{Script: Adlam}   (Short: \p{Sc=Adlm}) (88: U+1E900..1E94B,
                     U+1E950..1E959, U+1E95E..1E95F)
 \p{Script: Adlm}    \p{Script=Adlam} (88)
 \p{Script: Aghb}    \p{Script=Caucasian_Albanian} (=
                     \p{Script_Extensions=
                     Caucasian_Albanian}) (53)
 \p{Script: Ahom}    \p{Script_Extensions=Ahom} (Short: \p{Sc=
```

Ahom}, \p{Ahom}) (58)

\p{Script: Anatolian_Hieroglyphs} \p{Script_Extensions=
    Anatolian_Hieroglyphs} (Short: \p{Sc=
    Hluw}, \p{Hluw}) (583)

\p{Script: Arab} \p{Script=Arabic} (1291)

\p{Script: Arabic} (Short: \p{Sc=Arab}) (1291: U+0600..0604,
    U+0606..060B, U+060D..061A, U+061C,
    U+061E, U+0620..063F ...)

\p{Script: Armenian} \p{Script_Extensions=Armenian} (Short:
    \p{Sc=Armn}, \p{Armn}) (96)

\p{Script: Armi} \p{Script=Imperial_Aramaic} (=
    \p{Script_Extensions=Imperial_Aramaic})
    (31)

\p{Script: Armn} \p{Script=Armenian} (=
    \p{Script_Extensions=Armenian}) (96)

\p{Script: Avestan} \p{Script_Extensions=Avestan} (Short:
    \p{Sc=Avst}, \p{Avst}) (61)

\p{Script: Avst} \p{Script=Avestan} (=
    \p{Script_Extensions=Avestan}) (61)

\p{Script: Bali} \p{Script=Balinese} (=
    \p{Script_Extensions=Balinese}) (121)

\p{Script: Balinese} \p{Script_Extensions=Balinese} (Short:
    \p{Sc=Bali}, \p{Bali}) (121)

\p{Script: Bamu} \p{Script=Bamum} (= \p{Script_Extensions=
    Bamum}) (657)

\p{Script: Bamum} \p{Script_Extensions=Bamum} (Short: \p{Sc=
    Bamu}, \p{Bamu}) (657)

\p{Script: Bass} \p{Script=Bassa_Vah} (=
    \p{Script_Extensions=Bassa_Vah}) (36)

\p{Script: Bassa_Vah} \p{Script_Extensions=Bassa_Vah} (Short:
    \p{Sc=Bass}, \p{Bass}) (36)

\p{Script: Batak} \p{Script_Extensions=Batak} (Short: \p{Sc=
    Batk}, \p{Batk}) (56)

\p{Script: Batk} \p{Script=Batak} (= \p{Script_Extensions=
    Batak}) (56)

\p{Script: Beng} \p{Script=Bengali} (96)

\p{Script: Bengali} (Short: \p{Sc=Beng}) (96: U+0980..0983,
    U+0985..098C, U+098F..0990,
    U+0993..09A8, U+09AA..09B0, U+09B2 ...)

\p{Script: Bhaiksuki} \p{Script_Extensions=Bhaiksuki} (Short:

```
                            \p{Sc=Bhks}, \p{Bhks}) (97)
     \p{Script: Bhks}        \p{Script=Bhaiksuki} (=
                            \p{Script_Extensions=Bhaiksuki}) (97)
     \p{Script: Bopo}        \p{Script=Bopomofo} (77)
     \p{Script: Bopomofo}    (Short: \p{Sc=Bopo}) (77: U+02EA..02EB,
                            U+3105..312F, U+31A0..31BF)
     \p{Script: Brah}        \p{Script=Brahmi} (= \p{Script_Extensions=
                            Brahmi}) (109)
     \p{Script: Brahmi}      \p{Script_Extensions=Brahmi} (Short:
                            \p{Sc=Brah}, \p{Brah}) (109)
     \p{Script: Brai}        \p{Script=Braille} (=
                            \p{Script_Extensions=Braille}) (256)
     \p{Script: Braille}     \p{Script_Extensions=Braille} (Short:
                            \p{Sc=Brai}, \p{Brai}) (256)
     \p{Script: Bugi}        \p{Script=Buginese} (30)
     \p{Script: Buginese}    (Short: \p{Sc=Bugi}) (30: U+1A00..1A1B,
                            U+1A1E..1A1F)
     \p{Script: Buhd}        \p{Script=Buhid} (20)
     \p{Script: Buhid}       (Short: \p{Sc=Buhd}) (20: U+1740..1753)
     \p{Script: Cakm}        \p{Script=Chakma} (71)
     \p{Script: Canadian_Aboriginal} \p{Script_Extensions=
                            Canadian_Aboriginal} (Short: \p{Sc=
                            Cans}, \p{Cans}) (710)
     \p{Script: Cans}        \p{Script=Canadian_Aboriginal} (=
                            \p{Script_Extensions=
                            Canadian_Aboriginal}) (710)
     \p{Script: Cari}        \p{Script=Carian} (= \p{Script_Extensions=
                            Carian}) (49)
     \p{Script: Carian}      \p{Script_Extensions=Carian} (Short:
                            \p{Sc=Cari}, \p{Cari}) (49)
     \p{Script: Caucasian_Albanian} \p{Script_Extensions=
                            Caucasian_Albanian} (Short: \p{Sc=Aghb},
                            \p{Aghb}) (53)
     \p{Script: Chakma}      (Short: \p{Sc=Cakm}) (71: U+11100..11134,
                            U+11136..11147)
     \p{Script: Cham}        \p{Script_Extensions=Cham} (Short: \p{Sc=
                            Cham}, \p{Cham}) (83)
     \p{Script: Cher}        \p{Script=Cherokee} (=
                            \p{Script_Extensions=Cherokee}) (172)
     \p{Script: Cherokee}    \p{Script_Extensions=Cherokee} (Short:
```

```
                     \p{Sc=Cher}, \p{Cher}) (172)
\p{Script: Chorasmian}  \p{Script_Extensions=Chorasmian} (Short:
                     \p{Sc=Chrs}, \p{Chrs}) (28)
\p{Script: Chrs}      \p{Script=Chorasmian} (=
                     \p{Script_Extensions=Chorasmian}) (28)
\p{Script: Common}    (Short: \p{Sc=Zyyy}) (8087: [\x00-\x20!
                     \"#\$\%&\'\(\)*+,\-.\/0-9:;<=>?\@\[\\\]
                     \^_`\{\|\}~\x7f-\xa9\xab-\xb9\xbb-\xbf
                     \xd7\xf7], U+02B9..02DF, U+02E5..02E9,
                     U+02EC..02FF, U+0374, U+037E ...)
\p{Script: Copt}      \p{Script=Coptic} (137)
\p{Script: Coptic}    (Short: \p{Sc=Copt}) (137: U+03E2..03EF,
                     U+2C80..2CF3, U+2CF9..2CFF)
\p{Script: Cprt}      \p{Script=Cypriot} (55)
\p{Script: Cuneiform}  \p{Script_Extensions=Cuneiform} (Short:
                     \p{Sc=Xsux}, \p{Xsux}) (1234)
\p{Script: Cypriot}   (Short: \p{Sc=Cprt}) (55: U+10800..10805,
                     U+10808, U+1080A..10835, U+10837..10838,
                     U+1083C, U+1083F)
\p{Script: Cyrillic}  (Short: \p{Sc=Cyrl}) (443: U+0400..0484,
                     U+0487..052F, U+1C80..1C88, U+1D2B,
                     U+1D78, U+2DE0..2DFF ...)
\p{Script: Cyrl}      \p{Script=Cyrillic} (443)
\p{Script: Deseret}   \p{Script_Extensions=Deseret} (Short:
                     \p{Sc=Dsrt}, \p{Dsrt}) (80)
\p{Script: Deva}      \p{Script=Devanagari} (154)
\p{Script: Devanagari}  (Short: \p{Sc=Deva}) (154: U+0900..0950,
                     U+0955..0963, U+0966..097F, U+A8E0..A8FF)
\p{Script: Diak}      \p{Script=Dives_Akuru} (=
                     \p{Script_Extensions=Dives_Akuru}) (72)
\p{Script: Dives_Akuru}  \p{Script_Extensions=Dives_Akuru} (Short:
                     \p{Sc=Diak}, \p{Diak}) (72)
\p{Script: Dogr}      \p{Script=Dogra} (60)
\p{Script: Dogra}     (Short: \p{Sc=Dogr}) (60: U+11800..1183B)
\p{Script: Dsrt}      \p{Script=Deseret} (=
                     \p{Script_Extensions=Deseret}) (80)
\p{Script: Dupl}      \p{Script=Duployan} (143)
\p{Script: Duployan}  (Short: \p{Sc=Dupl}) (143: U+1BC00..1BC6A,
                     U+1BC70..1BC7C, U+1BC80..1BC88,
                     U+1BC90..1BC99, U+1BC9C..1BC9F)
```

\p{Script: Egyp}   \p{Script=Egyptian_Hieroglyphs} (=
                \p{Script_Extensions=
                Egyptian_Hieroglyphs}) (1080)
\p{Script: Egyptian_Hieroglyphs} \p{Script_Extensions=
                Egyptian_Hieroglyphs} (Short: \p{Sc=
                Egyp}, \p{Egyp}) (1080)
\p{Script: Elba}   \p{Script=Elbasan} (=
                \p{Script_Extensions=Elbasan}) (40)
\p{Script: Elbasan}   \p{Script_Extensions=Elbasan} (Short:
                \p{Sc=Elba}, \p{Elba}) (40)
\p{Script: Elym}   \p{Script=Elymaic} (=
                \p{Script_Extensions=Elymaic}) (23)
\p{Script: Elymaic}   \p{Script_Extensions=Elymaic} (Short:
                \p{Sc=Elym}, \p{Elym}) (23)
\p{Script: Ethi}   \p{Script=Ethiopic} (=
                \p{Script_Extensions=Ethiopic}) (495)
\p{Script: Ethiopic}   \p{Script_Extensions=Ethiopic} (Short:
                \p{Sc=Ethi}, \p{Ethi}) (495)
\p{Script: Geor}   \p{Script=Georgian} (173)
\p{Script: Georgian}   (Short: \p{Sc=Geor}) (173: U+10A0..10C5,
                U+10C7, U+10CD, U+10D0..10FA,
                U+10FC..10FF, U+1C90..1CBA ...)
\p{Script: Glag}   \p{Script=Glagolitic} (132)
\p{Script: Glagolitic} (Short: \p{Sc=Glag}) (132: U+2C00..2C2E,
                U+2C30..2C5E, U+1E000..1E006,
                U+1E008..1E018, U+1E01B..1E021,
                U+1E023..1E024 ...)
\p{Script: Gong}   \p{Script=Gunjala_Gondi} (63)
\p{Script: Gonm}   \p{Script=Masaram_Gondi} (75)
\p{Script: Goth}   \p{Script=Gothic} (= \p{Script_Extensions=
                Gothic}) (27)
\p{Script: Gothic}   \p{Script_Extensions=Gothic} (Short:
                \p{Sc=Goth}, \p{Goth}) (27)
\p{Script: Gran}   \p{Script=Grantha} (85)
\p{Script: Grantha}   (Short: \p{Sc=Gran}) (85: U+11300..11303,
                U+11305..1130C, U+1130F..11310,
                U+11313..11328, U+1132A..11330,
                U+11332..11333 ...)
\p{Script: Greek}   (Short: \p{Sc=Grek}) (518: U+0370..0373,
                U+0375..0377, U+037A..037D, U+037F,

```
                    U+0384, U+0386 ...)
    \p{Script: Grek}        \p{Script=Greek} (518)
    \p{Script: Gujarati}    (Short: \p{Sc=Gujr}) (91: U+0A81..0A83,
                    U+0A85..0A8D, U+0A8F..0A91,
                    U+0A93..0AA8, U+0AAA..0AB0, U+0AB2..0AB3
                    ...)
    \p{Script: Gujr}        \p{Script=Gujarati} (91)
    \p{Script: Gunjala_Gondi} (Short: \p{Sc=Gong}) (63:
                    U+11D60..11D65, U+11D67..11D68,
                    U+11D6A..11D8E, U+11D90..11D91,
                    U+11D93..11D98, U+11DA0..11DA9)
    \p{Script: Gurmukhi}    (Short: \p{Sc=Guru}) (80: U+0A01..0A03,
                    U+0A05..0A0A, U+0A0F..0A10,
                    U+0A13..0A28, U+0A2A..0A30, U+0A32..0A33
                    ...)
    \p{Script: Guru}        \p{Script=Gurmukhi} (80)
    \p{Script: Han}         (Short: \p{Sc=Han}) (94_204: U+2E80..2E99,
                    U+2E9B..2EF3, U+2F00..2FD5, U+3005,
                    U+3007, U+3021..3029 ...)
    \p{Script: Hang}        \p{Script=Hangul} (11_739)
    \p{Script: Hangul}      (Short: \p{Sc=Hang}) (11_739:
                    U+1100..11FF, U+302E..302F,
                    U+3131..318E, U+3200..321E,
                    U+3260..327E, U+A960..A97C ...)
    \p{Script: Hani}        \p{Script=Han} (94_204)
    \p{Script: Hanifi_Rohingya} (Short: \p{Sc=Rohg}) (50:
                    U+10D00..10D27, U+10D30..10D39)
    \p{Script: Hano}        \p{Script=Hanunoo} (21)
    \p{Script: Hanunoo}     (Short: \p{Sc=Hano}) (21: U+1720..1734)
    \p{Script: Hatr}        \p{Script=Hatran} (= \p{Script_Extensions=
                    Hatran}) (26)
    \p{Script: Hatran}      \p{Script_Extensions=Hatran} (Short:
                    \p{Sc=Hatr}, \p{Hatr}) (26)
    \p{Script: Hebr}        \p{Script=Hebrew} (= \p{Script_Extensions=
                    Hebrew}) (134)
    \p{Script: Hebrew}      \p{Script_Extensions=Hebrew} (Short:
                    \p{Sc=Hebr}, \p{Hebr}) (134)
    \p{Script: Hira}        \p{Script=Hiragana} (379)
    \p{Script: Hiragana}    (Short: \p{Sc=Hira}) (379: U+3041..3096,
                    U+309D..309F, U+1B001..1B11E,
```

U+1B150..1B152, U+1F200)

\p{Script: Hluw}   \p{Script=Anatolian_Hieroglyphs} (=
                \p{Script_Extensions=
                Anatolian_Hieroglyphs}) (583)

\p{Script: Hmng}   \p{Script=Pahawh_Hmong} (=
                \p{Script_Extensions=Pahawh_Hmong}) (127)

\p{Script: Hmnp}   \p{Script=Nyiakeng_Puachue_Hmong} (=
                \p{Script_Extensions=
                Nyiakeng_Puachue_Hmong}) (71)

\p{Script: Hung}   \p{Script=Old_Hungarian} (=
                \p{Script_Extensions=Old_Hungarian})
                (108)

\p{Script: Imperial_Aramaic} \p{Script_Extensions=
                Imperial_Aramaic} (Short: \p{Sc=Armi},
                \p{Armi}) (31)

\p{Script: Inherited}   (Short: \p{Sc=Zinh}) (573: U+0300..036F,
                U+0485..0486, U+064B..0655, U+0670,
                U+0951..0954, U+1AB0..1AC0 ...)

\p{Script: Inscriptional_Pahlavi} \p{Script_Extensions=
                Inscriptional_Pahlavi} (Short: \p{Sc=
                Phli}, \p{Phli}) (27)

\p{Script: Inscriptional_Parthian} \p{Script_Extensions=
                Inscriptional_Parthian} (Short: \p{Sc=
                Prti}, \p{Prti}) (30)

\p{Script: Ital}   \p{Script=Old_Italic} (=
                \p{Script_Extensions=Old_Italic}) (39)

\p{Script: Java}   \p{Script=Javanese} (90)

\p{Script: Javanese}   (Short: \p{Sc=Java}) (90: U+A980..A9CD,
                U+A9D0..A9D9, U+A9DE..A9DF)

\p{Script: Kaithi}   (Short: \p{Sc=Kthi}) (67: U+11080..110C1,
                U+110CD)

\p{Script: Kali}   \p{Script=Kayah_Li} (47)

\p{Script: Kana}   \p{Script=Katakana} (304)

\p{Script: Kannada}   (Short: \p{Sc=Knda}) (89: U+0C80..0C8C,
                U+0C8E..0C90, U+0C92..0CA8,
                U+0CAA..0CB3, U+0CB5..0CB9, U+0CBC..0CC4
                ...)

\p{Script: Katakana}   (Short: \p{Sc=Kana}) (304: U+30A1..30FA,
                U+30FD..30FF, U+31F0..31FF,
                U+32D0..32FE, U+3300..3357, U+FF66..FF6F

                         ...)
      \p{Script: Kayah_Li}   (Short: \p{Sc=Kali}) (47: U+A900..A92D,
                         U+A92F)
      \p{Script: Khar}   \p{Script=Kharoshthi} (=
                         \p{Script_Extensions=Kharoshthi}) (68)
      \p{Script: Kharoshthi} \p{Script_Extensions=Kharoshthi} (Short:
                         \p{Sc=Khar}, \p{Khar}) (68)
      \p{Script: Khitan_Small_Script} \p{Script_Extensions=
                         Khitan_Small_Script} (Short: \p{Sc=
                         Kits}, \p{Kits}) (471)
      \p{Script: Khmer}   \p{Script_Extensions=Khmer} (Short: \p{Sc=
                         Khmr}, \p{Khmr}) (146)
      \p{Script: Khmr}   \p{Script=Khmer} (= \p{Script_Extensions=
                         Khmer}) (146)
      \p{Script: Khoj}   \p{Script=Khojki} (62)
      \p{Script: Khojki}   (Short: \p{Sc=Khoj}) (62: U+11200..11211,
                         U+11213..1123E)
      \p{Script: Khudawadi}   (Short: \p{Sc=Sind}) (69: U+112B0..112EA,
                         U+112F0..112F9)
      \p{Script: Kits}   \p{Script=Khitan_Small_Script} (=
                         \p{Script_Extensions=
                         Khitan_Small_Script}) (471)
      \p{Script: Knda}   \p{Script=Kannada} (89)
      \p{Script: Kthi}   \p{Script=Kaithi} (67)
      \p{Script: Lana}   \p{Script=Tai_Tham} (=
                         \p{Script_Extensions=Tai_Tham}) (127)
      \p{Script: Lao}   \p{Script_Extensions=Lao} (Short: \p{Sc=
                         Lao}, \p{Lao}) (82)
      \p{Script: Laoo}   \p{Script=Lao} (= \p{Script_Extensions=
                         Lao}) (82)
      \p{Script: Latin}   (Short: \p{Sc=Latn}) (1374: [A-Za-z\xaa
                         \xba\xc0-\xd6\xd8-\xf6\xf8-\xff],
                         U+0100..02B8, U+02E0..02E4,
                         U+1D00..1D25, U+1D2C..1D5C, U+1D62..1D65
                         ...)
      \p{Script: Latn}   \p{Script=Latin} (1374)
      \p{Script: Lepc}   \p{Script=Lepcha} (= \p{Script_Extensions=
                         Lepcha}) (74)
      \p{Script: Lepcha}   \p{Script_Extensions=Lepcha} (Short:
                         \p{Sc=Lepc}, \p{Lepc}) (74)

```
\p{Script: Limb}        \p{Script=Limbu} (68)
\p{Script: Limbu}        (Short: \p{Sc=Limb}) (68: U+1900..191E,
                 U+1920..192B, U+1930..193B, U+1940,
                 U+1944..194F)
\p{Script: Lina}        \p{Script=Linear_A} (341)
\p{Script: Linb}        \p{Script=Linear_B} (211)
\p{Script: Linear_A}    (Short: \p{Sc=Lina}) (341: U+10600..10736,
                 U+10740..10755, U+10760..10767)
\p{Script: Linear_B}    (Short: \p{Sc=Linb}) (211: U+10000..1000B,
                 U+1000D..10026, U+10028..1003A,
                 U+1003C..1003D, U+1003F..1004D,
                 U+10050..1005D ...)
\p{Script: Lisu}        \p{Script_Extensions=Lisu} (Short: \p{Sc=
                 Lisu}, \p{Lisu}) (49)
\p{Script: Lyci}        \p{Script=Lycian} (= \p{Script_Extensions=
                 Lycian}) (29)
\p{Script: Lycian}      \p{Script_Extensions=Lycian} (Short:
                 \p{Sc=Lyci}, \p{Lyci}) (29)
\p{Script: Lydi}        \p{Script=Lydian} (= \p{Script_Extensions=
                 Lydian}) (27)
\p{Script: Lydian}      \p{Script_Extensions=Lydian} (Short:
                 \p{Sc=Lydi}, \p{Lydi}) (27)
\p{Script: Mahajani}    (Short: \p{Sc=Mahj}) (39: U+11150..11176)
\p{Script: Mahj}        \p{Script=Mahajani} (39)
\p{Script: Maka}        \p{Script=Makasar} (=
                 \p{Script_Extensions=Makasar}) (25)
\p{Script: Makasar}     \p{Script_Extensions=Makasar} (Short:
                 \p{Sc=Maka}, \p{Maka}) (25)
\p{Script: Malayalam}   (Short: \p{Sc=Mlym}) (118: U+0D00..0D0C,
                 U+0D0E..0D10, U+0D12..0D44,
                 U+0D46..0D48, U+0D4A..0D4F, U+0D54..0D63
                 ...)
\p{Script: Mand}        \p{Script=Mandaic} (29)
\p{Script: Mandaic}     (Short: \p{Sc=Mand}) (29: U+0840..085B,
                 U+085E)
\p{Script: Mani}        \p{Script=Manichaean} (51)
\p{Script: Manichaean}  (Short: \p{Sc=Mani}) (51: U+10AC0..10AE6,
                 U+10AEB..10AF6)
\p{Script: Marc}        \p{Script=Marchen} (=
                 \p{Script_Extensions=Marchen}) (68)
```

\p{Script: Marchen}   \p{Script_Extensions=Marchen} (Short:
                \p{Sc=Marc}, \p{Marc}) (68)
\p{Script: Masaram_Gondi} (Short: \p{Sc=Gonm}) (75:
                U+11D00..11D06, U+11D08..11D09,
                U+11D0B..11D36, U+11D3A, U+11D3C..11D3D,
                U+11D3F..11D47 ...)
\p{Script: Medefaidrin} \p{Script_Extensions=Medefaidrin} (Short:
                \p{Sc=Medf}, \p{Medf}) (91)
\p{Script: Medf}   \p{Script=Medefaidrin} (=
                \p{Script_Extensions=Medefaidrin}) (91)
\p{Script: Meetei_Mayek} \p{Script_Extensions=Meetei_Mayek}
                (Short: \p{Sc=Mtei}, \p{Mtei}) (79)
\p{Script: Mend}   \p{Script=Mende_Kikakui} (=
                \p{Script_Extensions=Mende_Kikakui})
                (213)
\p{Script: Mende_Kikakui} \p{Script_Extensions=Mende_Kikakui}
                (Short: \p{Sc=Mend}, \p{Mend}) (213)
\p{Script: Merc}   \p{Script=Meroitic_Cursive} (=
                \p{Script_Extensions=Meroitic_Cursive})
                (90)
\p{Script: Mero}   \p{Script=Meroitic_Hieroglyphs} (=
                \p{Script_Extensions=
                Meroitic_Hieroglyphs}) (32)
\p{Script: Meroitic_Cursive} \p{Script_Extensions=
                Meroitic_Cursive} (Short: \p{Sc=Merc},
                \p{Merc}) (90)
\p{Script: Meroitic_Hieroglyphs} \p{Script_Extensions=
                Meroitic_Hieroglyphs} (Short: \p{Sc=
                Mero}, \p{Mero}) (32)
\p{Script: Miao}   \p{Script_Extensions=Miao} (Short: \p{Sc=
                Miao}, \p{Miao}) (149)
\p{Script: Mlym}   \p{Script=Malayalam} (118)
\p{Script: Modi}   (Short: \p{Sc=Modi}) (79: U+11600..11644,
                U+11650..11659)
\p{Script: Mong}   \p{Script=Mongolian} (167)
\p{Script: Mongolian}   (Short: \p{Sc=Mong}) (167: U+1800..1801,
                U+1804, U+1806..180E, U+1810..1819,
                U+1820..1878, U+1880..18AA ...)
\p{Script: Mro}   \p{Script_Extensions=Mro} (Short: \p{Sc=
                Mro}, \p{Mro}) (43)

\p{Script: Mroo}   \p{Script=Mro} (= \p{Script_Extensions=
            Mro}) (43)
\p{Script: Mtei}   \p{Script=Meetei_Mayek} (=
            \p{Script_Extensions=Meetei_Mayek}) (79)
\p{Script: Mult}   \p{Script=Multani} (38)
\p{Script: Multani}   (Short: \p{Sc=Mult}) (38: U+11280..11286,
            U+11288, U+1128A..1128D, U+1128F..1129D,
            U+1129F..112A9)
\p{Script: Myanmar}   (Short: \p{Sc=Mymr}) (223: U+1000..109F,
            U+A9E0..A9FE, U+AA60..AA7F)
\p{Script: Mymr}   \p{Script=Myanmar} (223)
\p{Script: Nabataean} \p{Script_Extensions=Nabataean} (Short:
            \p{Sc=Nbat}, \p{Nbat}) (40)
\p{Script: Nand}   \p{Script=Nandinagari} (65)
\p{Script: Nandinagari} (Short: \p{Sc=Nand}) (65: U+119A0..119A7,
            U+119AA..119D7, U+119DA..119E4)
\p{Script: Narb}   \p{Script=Old_North_Arabian} (=
            \p{Script_Extensions=Old_North_Arabian})
            (32)
\p{Script: Nbat}   \p{Script=Nabataean} (=
            \p{Script_Extensions=Nabataean}) (40)
\p{Script: New_Tai_Lue} \p{Script_Extensions=New_Tai_Lue} (Short:
            \p{Sc=Talu}, \p{Talu}) (83)
\p{Script: Newa}   \p{Script_Extensions=Newa} (Short: \p{Sc=
            Newa}, \p{Newa}) (97)
\p{Script: Nko}   \p{Script_Extensions=Nko} (Short: \p{Sc=
            Nko}, \p{Nko}) (62)
\p{Script: Nkoo}   \p{Script=Nko} (= \p{Script_Extensions=
            Nko}) (62)
\p{Script: Nshu}   \p{Script=Nushu} (= \p{Script_Extensions=
            Nushu}) (397)
\p{Script: Nushu}   \p{Script_Extensions=Nushu} (Short: \p{Sc=
            Nshu}, \p{Nshu}) (397)
\p{Script: Nyiakeng_Puachue_Hmong} \p{Script_Extensions=
            Nyiakeng_Puachue_Hmong} (Short: \p{Sc=
            Hmnp}, \p{Hmnp}) (71)
\p{Script: Ogam}   \p{Script=Ogham} (= \p{Script_Extensions=
            Ogham}) (29)
\p{Script: Ogham}   \p{Script_Extensions=Ogham} (Short: \p{Sc=
            Ogam}, \p{Ogam}) (29)

\p{Script: Ol_Chiki}　\p{Script_Extensions=Ol_Chiki} (Short:
　　　　　　\p{Sc=Olck}, \p{Olck}) (48)
\p{Script: Olck}　　　\p{Script=Ol_Chiki} (=
　　　　　　\p{Script_Extensions=Ol_Chiki}) (48)
\p{Script: Old_Hungarian} \p{Script_Extensions=Old_Hungarian}
　　　　　　(Short: \p{Sc=Hung}, \p{Hung}) (108)
\p{Script: Old_Italic}　\p{Script_Extensions=Old_Italic} (Short:
　　　　　　\p{Sc=Ital}, \p{Ital}) (39)
\p{Script: Old_North_Arabian} \p{Script_Extensions=
　　　　　　Old_North_Arabian} (Short: \p{Sc=Narb},
　　　　　　\p{Narb}) (32)
\p{Script: Old_Permic}　(Short: \p{Sc=Perm}) (43: U+10350..1037A)
\p{Script: Old_Persian} \p{Script_Extensions=Old_Persian} (Short:
　　　　　　\p{Sc=Xpeo}, \p{Xpeo}) (50)
\p{Script: Old_Sogdian} \p{Script_Extensions=Old_Sogdian} (Short:
　　　　　　\p{Sc=Sogo}, \p{Sogo}) (40)
\p{Script: Old_South_Arabian} \p{Script_Extensions=
　　　　　　Old_South_Arabian} (Short: \p{Sc=Sarb},
　　　　　　\p{Sarb}) (32)
\p{Script: Old_Turkic}　\p{Script_Extensions=Old_Turkic} (Short:
　　　　　　\p{Sc=Orkh}, \p{Orkh}) (73)
\p{Script: Oriya}　　　(Short: \p{Sc=Orya}) (91: U+0B01..0B03,
　　　　　　U+0B05..0B0C, U+0B0F..0B10,
　　　　　　U+0B13..0B28, U+0B2A..0B30, U+0B32..0B33
　　　　　　...)
\p{Script: Orkh}　　　\p{Script=Old_Turkic} (=
　　　　　　\p{Script_Extensions=Old_Turkic}) (73)
\p{Script: Orya}　　　\p{Script=Oriya} (91)
\p{Script: Osage}　　　\p{Script_Extensions=Osage} (Short: \p{Sc=
　　　　　　Osge}, \p{Osge}) (72)
\p{Script: Osge}　　　\p{Script=Osage} (= \p{Script_Extensions=
　　　　　　Osage}) (72)
\p{Script: Osma}　　　\p{Script=Osmanya} (=
　　　　　　\p{Script_Extensions=Osmanya}) (40)
\p{Script: Osmanya}　　\p{Script_Extensions=Osmanya} (Short:
　　　　　　\p{Sc=Osma}, \p{Osma}) (40)
\p{Script: Pahawh_Hmong} \p{Script_Extensions=Pahawh_Hmong}
　　　　　　(Short: \p{Sc=Hmng}, \p{Hmng}) (127)
\p{Script: Palm}　　　\p{Script=Palmyrene} (=
　　　　　　\p{Script_Extensions=Palmyrene}) (32)

\p{Script: Palmyrene} \p{Script_Extensions=Palmyrene} (Short:
              \p{Sc=Palm}, \p{Palm}) (32)
\p{Script: Pau_Cin_Hau} \p{Script_Extensions=Pau_Cin_Hau} (Short:
              \p{Sc=Pauc}, \p{Pauc}) (57)
\p{Script: Pauc}   \p{Script=Pau_Cin_Hau} (=
              \p{Script_Extensions=Pau_Cin_Hau}) (57)
\p{Script: Perm}   \p{Script=Old_Permic} (43)
\p{Script: Phag}   \p{Script=Phags_Pa} (56)
\p{Script: Phags_Pa}   (Short: \p{Sc=Phag}) (56: U+A840..A877)
\p{Script: Phli}   \p{Script=Inscriptional_Pahlavi} (=
              \p{Script_Extensions=
              Inscriptional_Pahlavi}) (27)
\p{Script: Phlp}   \p{Script=Psalter_Pahlavi} (29)
\p{Script: Phnx}   \p{Script=Phoenician} (=
              \p{Script_Extensions=Phoenician}) (29)
\p{Script: Phoenician} \p{Script_Extensions=Phoenician} (Short:
              \p{Sc=Phnx}, \p{Phnx}) (29)
\p{Script: Plrd}   \p{Script=Miao} (= \p{Script_Extensions=
              Miao}) (149)
\p{Script: Prti}   \p{Script=Inscriptional_Parthian} (=
              \p{Script_Extensions=
              Inscriptional_Parthian}) (30)
\p{Script: Psalter_Pahlavi} (Short: \p{Sc=Phlp}) (29:
              U+10B80..10B91, U+10B99..10B9C,
              U+10BA9..10BAF)
\p{Script: Qaac}   \p{Script=Coptic} (137)
\p{Script: Qaai}   \p{Script=Inherited} (573)
\p{Script: Rejang}   \p{Script_Extensions=Rejang} (Short:
              \p{Sc=Rjng}, \p{Rjng}) (37)
\p{Script: Rjng}   \p{Script=Rejang} (= \p{Script_Extensions=
              Rejang}) (37)
\p{Script: Rohg}   \p{Script=Hanifi_Rohingya} (50)
\p{Script: Runic}   \p{Script_Extensions=Runic} (Short: \p{Sc=
              Runr}, \p{Runr}) (86)
\p{Script: Runr}   \p{Script=Runic} (= \p{Script_Extensions=
              Runic}) (86)
\p{Script: Samaritan}   \p{Script_Extensions=Samaritan} (Short:
              \p{Sc=Samr}, \p{Samr}) (61)
\p{Script: Samr}   \p{Script=Samaritan} (=
              \p{Script_Extensions=Samaritan}) (61)

\p{Script: Sarb}   \p{Script=Old_South_Arabian} (=
                \p{Script_Extensions=Old_South_Arabian})
                (32)
\p{Script: Saur}   \p{Script=Saurashtra} (=
                \p{Script_Extensions=Saurashtra}) (82)
\p{Script: Saurashtra} \p{Script_Extensions=Saurashtra} (Short:
                \p{Sc=Saur}, \p{Saur}) (82)
\p{Script: Sgnw}   \p{Script=SignWriting} (=
                \p{Script_Extensions=SignWriting}) (672)
\p{Script: Sharada}   (Short: \p{Sc=Shrd}) (96: U+11180..111DF)
\p{Script: Shavian}   \p{Script_Extensions=Shavian} (Short:
                \p{Sc=Shaw}, \p{Shaw}) (48)
\p{Script: Shaw}   \p{Script=Shavian} (=
                \p{Script_Extensions=Shavian}) (48)
\p{Script: Shrd}   \p{Script=Sharada} (96)
\p{Script: Sidd}   \p{Script=Siddham} (=
                \p{Script_Extensions=Siddham}) (92)
\p{Script: Siddham}   \p{Script_Extensions=Siddham} (Short:
                \p{Sc=Sidd}, \p{Sidd}) (92)
\p{Script: SignWriting} \p{Script_Extensions=SignWriting} (Short:
                \p{Sc=Sgnw}, \p{Sgnw}) (672)
\p{Script: Sind}   \p{Script=Khudawadi} (69)
\p{Script: Sinh}   \p{Script=Sinhala} (111)
\p{Script: Sinhala}   (Short: \p{Sc=Sinh}) (111: U+0D81..0D83,
                U+0D85..0D96, U+0D9A..0DB1,
                U+0DB3..0DBB, U+0DBD, U+0DC0..0DC6 ...)
\p{Script: Sogd}   \p{Script=Sogdian} (42)
\p{Script: Sogdian}   (Short: \p{Sc=Sogd}) (42: U+10F30..10F59)
\p{Script: Sogo}   \p{Script=Old_Sogdian} (=
                \p{Script_Extensions=Old_Sogdian}) (40)
\p{Script: Sora}   \p{Script=Sora_Sompeng} (=
                \p{Script_Extensions=Sora_Sompeng}) (35)
\p{Script: Sora_Sompeng} \p{Script_Extensions=Sora_Sompeng}
                (Short: \p{Sc=Sora}, \p{Sora}) (35)
\p{Script: Soyo}   \p{Script=Soyombo} (=
                \p{Script_Extensions=Soyombo}) (83)
\p{Script: Soyombo}   \p{Script_Extensions=Soyombo} (Short:
                \p{Sc=Soyo}, \p{Soyo}) (83)
\p{Script: Sund}   \p{Script=Sundanese} (=
                \p{Script_Extensions=Sundanese}) (72)

\p{Script: Sundanese}   \p{Script_Extensions=Sundanese} (Short:
              \p{Sc=Sund}, \p{Sund}) (72)
\p{Script: Sylo}   \p{Script=Syloti_Nagri} (45)
\p{Script: Syloti_Nagri} (Short: \p{Sc=Sylo}) (45: U+A800..A82C)
\p{Script: Syrc}   \p{Script=Syriac} (88)
\p{Script: Syriac}   (Short: \p{Sc=Syrc}) (88: U+0700..070D,
              U+070F..074A, U+074D..074F, U+0860..086A)
\p{Script: Tagalog}   (Short: \p{Sc=Tglg}) (20: U+1700..170C,
              U+170E..1714)
\p{Script: Tagb}   \p{Script=Tagbanwa} (18)
\p{Script: Tagbanwa}   (Short: \p{Sc=Tagb}) (18: U+1760..176C,
              U+176E..1770, U+1772..1773)
\p{Script: Tai_Le}   (Short: \p{Sc=Tale}) (35: U+1950..196D,
              U+1970..1974)
\p{Script: Tai_Tham}   \p{Script_Extensions=Tai_Tham} (Short:
              \p{Sc=Lana}, \p{Lana}) (127)
\p{Script: Tai_Viet}   \p{Script_Extensions=Tai_Viet} (Short:
              \p{Sc=Tavt}, \p{Tavt}) (72)
\p{Script: Takr}   \p{Script=Takri} (67)
\p{Script: Takri}   (Short: \p{Sc=Takr}) (67: U+11680..116B8,
              U+116C0..116C9)
\p{Script: Tale}   \p{Script=Tai_Le} (35)
\p{Script: Talu}   \p{Script=New_Tai_Lue} (=
              \p{Script_Extensions=New_Tai_Lue}) (83)
\p{Script: Tamil}   (Short: \p{Sc=Taml}) (123: U+0B82..0B83,
              U+0B85..0B8A, U+0B8E..0B90,
              U+0B92..0B95, U+0B99..0B9A, U+0B9C ...)
\p{Script: Taml}   \p{Script=Tamil} (123)
\p{Script: Tang}   \p{Script=Tangut} (= \p{Script_Extensions=
              Tangut}) (6914)
\p{Script: Tangut}   \p{Script_Extensions=Tangut} (Short:
              \p{Sc=Tang}, \p{Tang}) (6914)
\p{Script: Tavt}   \p{Script=Tai_Viet} (=
              \p{Script_Extensions=Tai_Viet}) (72)
\p{Script: Telu}   \p{Script=Telugu} (98)
\p{Script: Telugu}   (Short: \p{Sc=Telu}) (98: U+0C00..0C0C,
              U+0C0E..0C10, U+0C12..0C28,
              U+0C2A..0C39, U+0C3D..0C44, U+0C46..0C48
              ...)
\p{Script: Tfng}   \p{Script=Tifinagh} (=

```
                        \p{Script_Extensions=Tifinagh}) (59)
        \p{Script: Tglg}      \p{Script=Tagalog} (20)
        \p{Script: Thaa}      \p{Script=Thaana} (50)
        \p{Script: Thaana}    (Short: \p{Sc=Thaa}) (50: U+0780..07B1)
        \p{Script: Thai}      \p{Script_Extensions=Thai} (Short: \p{Sc=
                        Thai}, \p{Thai}) (86)
        \p{Script: Tibetan}   \p{Script_Extensions=Tibetan} (Short:
                        \p{Sc=Tibt}, \p{Tibt}) (207)
        \p{Script: Tibt}      \p{Script=Tibetan} (=
                        \p{Script_Extensions=Tibetan}) (207)
        \p{Script: Tifinagh}  \p{Script_Extensions=Tifinagh} (Short:
                        \p{Sc=Tfng}, \p{Tfng}) (59)
        \p{Script: Tirh}      \p{Script=Tirhuta} (82)
        \p{Script: Tirhuta}   (Short: \p{Sc=Tirh}) (82: U+11480..114C7,
                        U+114D0..114D9)
        \p{Script: Ugar}      \p{Script=Ugaritic} (=
                        \p{Script_Extensions=Ugaritic}) (31)
        \p{Script: Ugaritic}  \p{Script_Extensions=Ugaritic} (Short:
                        \p{Sc=Ugar}, \p{Ugar}) (31)
        \p{Script: Unknown}   \p{Script_Extensions=Unknown} (Short:
                        \p{Sc=Zzzz}, \p{Zzzz}) (970_188 plus all
                        above-Unicode code points)
        \p{Script: Vai}       \p{Script_Extensions=Vai} (Short: \p{Sc=
                        Vai}, \p{Vai}) (300)
        \p{Script: Vaii}      \p{Script=Vai} (= \p{Script_Extensions=
                        Vai}) (300)
        \p{Script: Wancho}    \p{Script_Extensions=Wancho} (Short:
                        \p{Sc=Wcho}, \p{Wcho}) (59)
        \p{Script: Wara}      \p{Script=Warang_Citi} (=
                        \p{Script_Extensions=Warang_Citi}) (84)
        \p{Script: Warang_Citi} \p{Script_Extensions=Warang_Citi} (Short:
                        \p{Sc=Wara}, \p{Wara}) (84)
        \p{Script: Wcho}      \p{Script=Wancho} (= \p{Script_Extensions=
                        Wancho}) (59)
        \p{Script: Xpeo}      \p{Script=Old_Persian} (=
                        \p{Script_Extensions=Old_Persian}) (50)
        \p{Script: Xsux}      \p{Script=Cuneiform} (=
                        \p{Script_Extensions=Cuneiform}) (1234)
        \p{Script: Yezi}      \p{Script=Yezidi} (47)
        \p{Script: Yezidi}    (Short: \p{Sc=Yezi}) (47: U+10E80..10EA9,
```

```
                    U+10EAB..10EAD, U+10EB0..10EB1)
     \p{Script: Yi}        (Short: \p{Sc=Yi}) (1220: U+A000..A48C,
                    U+A490..A4C6)
     \p{Script: Yiii}      \p{Script=Yi} (1220)
     \p{Script: Zanabazar_Square} \p{Script_Extensions=
                    Zanabazar_Square} (Short: \p{Sc=Zanb},
                    \p{Zanb}) (72)
     \p{Script: Zanb}      \p{Script=Zanabazar_Square} (=
                    \p{Script_Extensions=Zanabazar_Square})
                    (72)
     \p{Script: Zinh}      \p{Script=Inherited} (573)
     \p{Script: Zyyy}      \p{Script=Common} (8087)
     \p{Script: Zzzz}      \p{Script=Unknown} (=
                    \p{Script_Extensions=Unknown}) (970_188
                    plus all above-Unicode code points)
     \p{Script_Extensions: Adlam} (Short: \p{Scx=Adlm}, \p{Adlm}) (89:
                    U+0640, U+1E900..1E94B, U+1E950..1E959,
                    U+1E95E..1E95F)
     \p{Script_Extensions: Adlm} \p{Script_Extensions=Adlam} (89)
     \p{Script_Extensions: Aghb} \p{Script_Extensions=
                    Caucasian_Albanian} (53)
     \p{Script_Extensions: Ahom} (Short: \p{Scx=Ahom}, \p{Ahom}) (58:
                    U+11700..1171A, U+1171D..1172B,
                    U+11730..1173F)
     \p{Script_Extensions: Anatolian_Hieroglyphs} (Short: \p{Scx=Hluw},
                    \p{Hluw}) (583: U+14400..14646)
     \p{Script_Extensions: Arab} \p{Script_Extensions=Arabic} (1335)
     \p{Script_Extensions: Arabic} (Short: \p{Scx=Arab}, \p{Arab})
                    (1335: U+0600..0604, U+0606..061C,
                    U+061E..06DC, U+06DE..06FF,
                    U+0750..077F, U+08A0..08B4 ...)
     \p{Script_Extensions: Armenian} (Short: \p{Scx=Armn}, \p{Armn})
                    (96: U+0531..0556, U+0559..058A,
                    U+058D..058F, U+FB13..FB17)
     \p{Script_Extensions: Armi} \p{Script_Extensions=Imperial_Aramaic}
                    (31)
     \p{Script_Extensions: Armn} \p{Script_Extensions=Armenian} (96)
     \p{Script_Extensions: Avestan} (Short: \p{Scx=Avst}, \p{Avst})
                    (61: U+10B00..10B35, U+10B39..10B3F)
     \p{Script_Extensions: Avst} \p{Script_Extensions=Avestan} (61)
```

\p{Script_Extensions: Bali} \p{Script_Extensions=Balinese} (121)
\p{Script_Extensions: Balinese} (Short: \p{Scx=Bali}, \p{Bali})
                 (121: U+1B00..1B4B, U+1B50..1B7C)
\p{Script_Extensions: Bamu} \p{Script_Extensions=Bamum} (657)
\p{Script_Extensions: Bamum} (Short: \p{Scx=Bamu}, \p{Bamu}) (657:
                 U+A6A0..A6F7, U+16800..16A38)
\p{Script_Extensions: Bass} \p{Script_Extensions=Bassa_Vah} (36)
\p{Script_Extensions: Bassa_Vah} (Short: \p{Scx=Bass}, \p{Bass})
                 (36: U+16AD0..16AED, U+16AF0..16AF5)
\p{Script_Extensions: Batak} (Short: \p{Scx=Batk}, \p{Batk}) (56:
                 U+1BC0..1BF3, U+1BFC..1BFF)
\p{Script_Extensions: Batk} \p{Script_Extensions=Batak} (56)
\p{Script_Extensions: Beng} \p{Script_Extensions=Bengali} (113)
\p{Script_Extensions: Bengali} (Short: \p{Scx=Beng}, \p{Beng})
                 (113: U+0951..0952, U+0964..0965,
                 U+0980..0983, U+0985..098C,
                 U+098F..0990, U+0993..09A8 ...)
\p{Script_Extensions: Bhaiksuki} (Short: \p{Scx=Bhks}, \p{Bhks})
                 (97: U+11C00..11C08, U+11C0A..11C36,
                 U+11C38..11C45, U+11C50..11C6C)
\p{Script_Extensions: Bhks} \p{Script_Extensions=Bhaiksuki} (97)
\p{Script_Extensions: Bopo} \p{Script_Extensions=Bopomofo} (117)
\p{Script_Extensions: Bopomofo} (Short: \p{Scx=Bopo}, \p{Bopo})
                 (117: U+02EA..02EB, U+3001..3003,
                 U+3008..3011, U+3013..301F,
                 U+302A..302D, U+3030 ...)
\p{Script_Extensions: Brah} \p{Script_Extensions=Brahmi} (109)
\p{Script_Extensions: Brahmi} (Short: \p{Scx=Brah}, \p{Brah})
                 (109: U+11000..1104D, U+11052..1106F,
                 U+1107F)
\p{Script_Extensions: Brai} \p{Script_Extensions=Braille} (256)
\p{Script_Extensions: Braille} (Short: \p{Scx=Brai}, \p{Brai})
                 (256: U+2800..28FF)
\p{Script_Extensions: Bugi} \p{Script_Extensions=Buginese} (31)
\p{Script_Extensions: Buginese} (Short: \p{Scx=Bugi}, \p{Bugi})
                 (31: U+1A00..1A1B, U+1A1E..1A1F, U+A9CF)
\p{Script_Extensions: Buhd} \p{Script_Extensions=Buhid} (22)
\p{Script_Extensions: Buhid} (Short: \p{Scx=Buhd}, \p{Buhd}) (22:
                 U+1735..1736, U+1740..1753)
\p{Script_Extensions: Cakm} \p{Script_Extensions=Chakma} (91)

\p{Script_Extensions: Canadian_Aboriginal} (Short: \p{Scx=Cans},
                \p{Cans}) (710: U+1400..167F,
                U+18B0..18F5)
\p{Script_Extensions: Cans} \p{Script_Extensions=
                Canadian_Aboriginal} (710)
\p{Script_Extensions: Cari} \p{Script_Extensions=Carian} (49)
\p{Script_Extensions: Carian} (Short: \p{Scx=Cari}, \p{Cari}) (49:
                U+102A0..102D0)
\p{Script_Extensions: Caucasian_Albanian} (Short: \p{Scx=Aghb},
                \p{Aghb}) (53: U+10530..10563, U+1056F)
\p{Script_Extensions: Chakma} (Short: \p{Scx=Cakm}, \p{Cakm}) (91:
                U+09E6..09EF, U+1040..1049,
                U+11100..11134, U+11136..11147)
\p{Script_Extensions: Cham} (Short: \p{Scx=Cham}, \p{Cham}) (83:
                U+AA00..AA36, U+AA40..AA4D,
                U+AA50..AA59, U+AA5C..AA5F)
\p{Script_Extensions: Cher} \p{Script_Extensions=Cherokee} (172)
\p{Script_Extensions: Cherokee} (Short: \p{Scx=Cher}, \p{Cher})
                (172: U+13A0..13F5, U+13F8..13FD,
                U+AB70..ABBF)
\p{Script_Extensions: Chorasmian} (Short: \p{Scx=Chrs}, \p{Chrs})
                (28: U+10FB0..10FCB)
\p{Script_Extensions: Chrs} \p{Script_Extensions=Chorasmian} (28)
\p{Script_Extensions: Common} (Short: \p{Scx=Zyyy}, \p{Zyyy})
                (7661: [\x00-\x20!\"#\$\%&\'\(\)*+,\-.
                \/0-9:;<=>?\@\[\\\]\^_'\{\|\}~\x7f-\xa9
                \xab-\xb9\xbb-\xbf\xd7\xf7],
                U+02B9..02DF, U+02E5..02E9,
                U+02EC..02FF, U+0374, U+037E ...)
\p{Script_Extensions: Copt} \p{Script_Extensions=Coptic} (165)
\p{Script_Extensions: Coptic} (Short: \p{Scx=Copt}, \p{Copt})
                (165: U+03E2..03EF, U+2C80..2CF3,
                U+2CF9..2CFF, U+102E0..102FB)
\p{Script_Extensions: Cprt} \p{Script_Extensions=Cypriot} (112)
\p{Script_Extensions: Cuneiform} (Short: \p{Scx=Xsux}, \p{Xsux})
                (1234: U+12000..12399, U+12400..1246E,
                U+12470..12474, U+12480..12543)
\p{Script_Extensions: Cypriot} (Short: \p{Scx=Cprt}, \p{Cprt})
                (112: U+10100..10102, U+10107..10133,
                U+10137..1013F, U+10800..10805, U+10808,

U+1080A..10835 ...)
\p{Script_Extensions: Cyrillic} (Short: \p{Scx=Cyrl}, \p{Cyrl})
          (447: U+0400..052F, U+1C80..1C88,
          U+1D2B, U+1D78, U+1DF8, U+2DE0..2DFF ...)
\p{Script_Extensions: Cyrl} \p{Script_Extensions=Cyrillic} (447)
\p{Script_Extensions: Deseret} (Short: \p{Scx=Dsrt}, \p{Dsrt})
          (80: U+10400..1044F)
\p{Script_Extensions: Deva} \p{Script_Extensions=Devanagari} (210)
\p{Script_Extensions: Devanagari} (Short: \p{Scx=Deva}, \p{Deva})
          (210: U+0900..0952, U+0955..097F,
          U+1CD0..1CF6, U+1CF8..1CF9, U+20F0,
          U+A830..A839 ...)
\p{Script_Extensions: Diak} \p{Script_Extensions=Dives_Akuru} (72)
\p{Script_Extensions: Dives_Akuru} (Short: \p{Scx=Diak}, \p{Diak})
          (72: U+11900..11906, U+11909,
          U+1190C..11913, U+11915..11916,
          U+11918..11935, U+11937..11938 ...)
\p{Script_Extensions: Dogr} \p{Script_Extensions=Dogra} (82)
\p{Script_Extensions: Dogra} (Short: \p{Scx=Dogr}, \p{Dogr}) (82:
          U+0964..096F, U+A830..A839,
          U+11800..1183B)
\p{Script_Extensions: Dsrt} \p{Script_Extensions=Deseret} (80)
\p{Script_Extensions: Dupl} \p{Script_Extensions=Duployan} (147)
\p{Script_Extensions: Duployan} (Short: \p{Scx=Dupl}, \p{Dupl})
          (147: U+1BC00..1BC6A, U+1BC70..1BC7C,
          U+1BC80..1BC88, U+1BC90..1BC99,
          U+1BC9C..1BCA3)
\p{Script_Extensions: Egyp} \p{Script_Extensions=
          Egyptian_Hieroglyphs} (1080)
\p{Script_Extensions: Egyptian_Hieroglyphs} (Short: \p{Scx=Egyp},
          \p{Egyp}) (1080: U+13000..1342E,
          U+13430..13438)
\p{Script_Extensions: Elba} \p{Script_Extensions=Elbasan} (40)
\p{Script_Extensions: Elbasan} (Short: \p{Scx=Elba}, \p{Elba})
          (40: U+10500..10527)
\p{Script_Extensions: Elym} \p{Script_Extensions=Elymaic} (23)
\p{Script_Extensions: Elymaic} (Short: \p{Scx=Elym}, \p{Elym})
          (23: U+10FE0..10FF6)
\p{Script_Extensions: Ethi} \p{Script_Extensions=Ethiopic} (495)
\p{Script_Extensions: Ethiopic} (Short: \p{Scx=Ethi}, \p{Ethi})

                (495: U+1200..1248, U+124A..124D,
                U+1250..1256, U+1258, U+125A..125D,
                U+1260..1288 ...)
\p{Script_Extensions: Geor} \p{Script_Extensions=Georgian} (174)
\p{Script_Extensions: Georgian} (Short: \p{Scx=Geor}, \p{Geor})
                (174: U+10A0..10C5, U+10C7, U+10CD,
                U+10D0..10FF, U+1C90..1CBA, U+1CBD..1CBF
                ...)
\p{Script_Extensions: Glag} \p{Script_Extensions=Glagolitic} (136)
\p{Script_Extensions: Glagolitic} (Short: \p{Scx=Glag}, \p{Glag})
                (136: U+0484, U+0487, U+2C00..2C2E,
                U+2C30..2C5E, U+2E43, U+A66F ...)
\p{Script_Extensions: Gong} \p{Script_Extensions=Gunjala_Gondi}
                (65)
\p{Script_Extensions: Gonm} \p{Script_Extensions=Masaram_Gondi}
                (77)
\p{Script_Extensions: Goth} \p{Script_Extensions=Gothic} (27)
\p{Script_Extensions: Gothic} (Short: \p{Scx=Goth}, \p{Goth}) (27:
                U+10330..1034A)
\p{Script_Extensions: Gran} \p{Script_Extensions=Grantha} (116)
\p{Script_Extensions: Grantha} (Short: \p{Scx=Gran}, \p{Gran})
                (116: U+0951..0952, U+0964..0965,
                U+0BE6..0BF3, U+1CD0, U+1CD2..1CD3,
                U+1CF2..1CF4 ...)
\p{Script_Extensions: Greek} (Short: \p{Scx=Grek}, \p{Grek}) (522:
                U+0342, U+0345, U+0370..0373,
                U+0375..0377, U+037A..037D, U+037F ...)
\p{Script_Extensions: Grek} \p{Script_Extensions=Greek} (522)
\p{Script_Extensions: Gujarati} (Short: \p{Scx=Gujr}, \p{Gujr})
                (105: U+0951..0952, U+0964..0965,
                U+0A81..0A83, U+0A85..0A8D,
                U+0A8F..0A91, U+0A93..0AA8 ...)
\p{Script_Extensions: Gujr} \p{Script_Extensions=Gujarati} (105)
\p{Script_Extensions: Gunjala_Gondi} (Short: \p{Scx=Gong},
                \p{Gong}) (65: U+0964..0965,
                U+11D60..11D65, U+11D67..11D68,
                U+11D6A..11D8E, U+11D90..11D91,
                U+11D93..11D98 ...)
\p{Script_Extensions: Gurmukhi} (Short: \p{Scx=Guru}, \p{Guru})
                (94: U+0951..0952, U+0964..0965,

U+0A01..0A03, U+0A05..0A0A,
U+0A0F..0A10, U+0A13..0A28 ...)
\p{Script_Extensions: Guru} \p{Script_Extensions=Gurmukhi} (94)
\p{Script_Extensions: Han} (Short: \p{Scx=Han}, \p{Han}) (94_492:
U+2E80..2E99, U+2E9B..2EF3,
U+2F00..2FD5, U+3001..3003,
U+3005..3011, U+3013..301F ...)
\p{Script_Extensions: Hang} \p{Script_Extensions=Hangul} (11_775)
\p{Script_Extensions: Hangul} (Short: \p{Scx=Hang}, \p{Hang})
(11_775: U+1100..11FF, U+3001..3003,
U+3008..3011, U+3013..301F,
U+302E..3030, U+3037 ...)
\p{Script_Extensions: Hani} \p{Script_Extensions=Han} (94_492)
\p{Script_Extensions: Hanifi_Rohingya} (Short: \p{Scx=Rohg},
\p{Rohg}) (55: U+060C, U+061B, U+061F,
U+0640, U+06D4, U+10D00..10D27 ...)
\p{Script_Extensions: Hano} \p{Script_Extensions=Hanunoo} (23)
\p{Script_Extensions: Hanunoo} (Short: \p{Scx=Hano}, \p{Hano})
(23: U+1720..1736)
\p{Script_Extensions: Hatr} \p{Script_Extensions=Hatran} (26)
\p{Script_Extensions: Hatran} (Short: \p{Scx=Hatr}, \p{Hatr}) (26:
U+108E0..108F2, U+108F4..108F5,
U+108FB..108FF)
\p{Script_Extensions: Hebr} \p{Script_Extensions=Hebrew} (134)
\p{Script_Extensions: Hebrew} (Short: \p{Scx=Hebr}, \p{Hebr})
(134: U+0591..05C7, U+05D0..05EA,
U+05EF..05F4, U+FB1D..FB36,
U+FB38..FB3C, U+FB3E ...)
\p{Script_Extensions: Hira} \p{Script_Extensions=Hiragana} (431)
\p{Script_Extensions: Hiragana} (Short: \p{Scx=Hira}, \p{Hira})
(431: U+3001..3003, U+3008..3011,
U+3013..301F, U+3030..3035, U+3037,
U+303C..303D ...)
\p{Script_Extensions: Hluw} \p{Script_Extensions=
Anatolian_Hieroglyphs} (583)
\p{Script_Extensions: Hmng} \p{Script_Extensions=Pahawh_Hmong}
(127)
\p{Script_Extensions: Hmnp} \p{Script_Extensions=
Nyiakeng_Puachue_Hmong} (71)
\p{Script_Extensions: Hung} \p{Script_Extensions=Old_Hungarian}

(108)
\p{Script_Extensions: Imperial_Aramaic} (Short: \p{Scx=Armi},
          \p{Armi}) (31: U+10840..10855,
          U+10857..1085F)
\p{Script_Extensions: Inherited} (Short: \p{Scx=Zinh}, \p{Zinh})
          (503: U+0300..0341, U+0343..0344,
          U+0346..0362, U+0953..0954,
          U+1AB0..1AC0, U+1DC2..1DF7 ...)
\p{Script_Extensions: Inscriptional_Pahlavi} (Short: \p{Scx=Phli},
          \p{Phli}) (27: U+10B60..10B72,
          U+10B78..10B7F)
\p{Script_Extensions: Inscriptional_Parthian} (Short: \p{Scx=
          Prti}, \p{Prti}) (30: U+10B40..10B55,
          U+10B58..10B5F)
\p{Script_Extensions: Ital} \p{Script_Extensions=Old_Italic} (39)
\p{Script_Extensions: Java} \p{Script_Extensions=Javanese} (91)
\p{Script_Extensions: Javanese} (Short: \p{Scx=Java}, \p{Java})
          (91: U+A980..A9CD, U+A9CF..A9D9,
          U+A9DE..A9DF)
\p{Script_Extensions: Kaithi} (Short: \p{Scx=Kthi}, \p{Kthi}) (87:
          U+0966..096F, U+A830..A839,
          U+11080..110C1, U+110CD)
\p{Script_Extensions: Kali} \p{Script_Extensions=Kayah_Li} (48)
\p{Script_Extensions: Kana} \p{Script_Extensions=Katakana} (356)
\p{Script_Extensions: Kannada} (Short: \p{Scx=Knda}, \p{Knda})
          (104: U+0951..0952, U+0964..0965,
          U+0C80..0C8C, U+0C8E..0C90,
          U+0C92..0CA8, U+0CAA..0CB3 ...)
\p{Script_Extensions: Katakana} (Short: \p{Scx=Kana}, \p{Kana})
          (356: U+3001..3003, U+3008..3011,
          U+3013..301F, U+3030..3035, U+3037,
          U+303C..303D ...)
\p{Script_Extensions: Kayah_Li} (Short: \p{Scx=Kali}, \p{Kali})
          (48: U+A900..A92F)
\p{Script_Extensions: Khar} \p{Script_Extensions=Kharoshthi} (68)
\p{Script_Extensions: Kharoshthi} (Short: \p{Scx=Khar}, \p{Khar})
          (68: U+10A00..10A03, U+10A05..10A06,
          U+10A0C..10A13, U+10A15..10A17,
          U+10A19..10A35, U+10A38..10A3A ...)
\p{Script_Extensions: Khitan_Small_Script} (Short: \p{Scx=Kits},

\p{Kits}) (471: U+16FE4, U+18B00..18CD5)

\p{Script_Extensions: Khmer} (Short: \p{Scx=Khmr}, \p{Khmr}) (146:
                U+1780..17DD, U+17E0..17E9,
                U+17F0..17F9, U+19E0..19FF)

\p{Script_Extensions: Khmr} \p{Script_Extensions=Khmer} (146)

\p{Script_Extensions: Khoj} \p{Script_Extensions=Khojki} (82)

\p{Script_Extensions: Khojki} (Short: \p{Scx=Khoj}, \p{Khoj}) (82:
                U+0AE6..0AEF, U+A830..A839,
                U+11200..11211, U+11213..1123E)

\p{Script_Extensions: Khudawadi} (Short: \p{Scx=Sind}, \p{Sind})
                (81: U+0964..0965, U+A830..A839,
                U+112B0..112EA, U+112F0..112F9)

\p{Script_Extensions: Kits} \p{Script_Extensions=
                Khitan_Small_Script} (471)

\p{Script_Extensions: Knda} \p{Script_Extensions=Kannada} (104)

\p{Script_Extensions: Kthi} \p{Script_Extensions=Kaithi} (87)

\p{Script_Extensions: Lana} \p{Script_Extensions=Tai_Tham} (127)

\p{Script_Extensions: Lao} (Short: \p{Scx=Lao}, \p{Lao}) (82:
                U+0E81..0E82, U+0E84, U+0E86..0E8A,
                U+0E8C..0EA3, U+0EA5, U+0EA7..0EBD ...)

\p{Script_Extensions: Laoo} \p{Script_Extensions=Lao} (82)

\p{Script_Extensions: Latin} (Short: \p{Scx=Latn}, \p{Latn})
                (1403: [A-Za-z\xaa\xba\xc0-\xd6\xd8-
                \xf6\xf8-\xff], U+0100..02B8,
                U+02E0..02E4, U+0363..036F,
                U+0485..0486, U+0951..0952 ...)

\p{Script_Extensions: Latn} \p{Script_Extensions=Latin} (1403)

\p{Script_Extensions: Lepc} \p{Script_Extensions=Lepcha} (74)

\p{Script_Extensions: Lepcha} (Short: \p{Scx=Lepc}, \p{Lepc}) (74:
                U+1C00..1C37, U+1C3B..1C49, U+1C4D..1C4F)

\p{Script_Extensions: Limb} \p{Script_Extensions=Limbu} (69)

\p{Script_Extensions: Limbu} (Short: \p{Scx=Limb}, \p{Limb}) (69:
                U+0965, U+1900..191E, U+1920..192B,
                U+1930..193B, U+1940, U+1944..194F)

\p{Script_Extensions: Lina} \p{Script_Extensions=Linear_A} (386)

\p{Script_Extensions: Linb} \p{Script_Extensions=Linear_B} (268)

\p{Script_Extensions: Linear_A} (Short: \p{Scx=Lina}, \p{Lina})
                (386: U+10107..10133, U+10600..10736,
                U+10740..10755, U+10760..10767)

\p{Script_Extensions: Linear_B} (Short: \p{Scx=Linb}, \p{Linb})

        (268: U+10000..1000B, U+1000D..10026,
        U+10028..1003A, U+1003C..1003D,
        U+1003F..1004D, U+10050..1005D ...)

\p{Script_Extensions: Lisu} (Short: \p{Scx=Lisu}, \p{Lisu}) (49:
        U+A4D0..A4FF, U+11FB0)

\p{Script_Extensions: Lyci} \p{Script_Extensions=Lycian} (29)

\p{Script_Extensions: Lycian} (Short: \p{Scx=Lyci}, \p{Lyci}) (29:
        U+10280..1029C)

\p{Script_Extensions: Lydi} \p{Script_Extensions=Lydian} (27)

\p{Script_Extensions: Lydian} (Short: \p{Scx=Lydi}, \p{Lydi}) (27:
        U+10920..10939, U+1093F)

\p{Script_Extensions: Mahajani} (Short: \p{Scx=Mahj}, \p{Mahj})
        (61: U+0964..096F, U+A830..A839,
        U+11150..11176)

\p{Script_Extensions: Mahj} \p{Script_Extensions=Mahajani} (61)

\p{Script_Extensions: Maka} \p{Script_Extensions=Makasar} (25)

\p{Script_Extensions: Makasar} (Short: \p{Scx=Maka}, \p{Maka})
        (25: U+11EE0..11EF8)

\p{Script_Extensions: Malayalam} (Short: \p{Scx=Mlym}, \p{Mlym})
        (126: U+0951..0952, U+0964..0965,
        U+0D00..0D0C, U+0D0E..0D10,
        U+0D12..0D44, U+0D46..0D48 ...)

\p{Script_Extensions: Mand} \p{Script_Extensions=Mandaic} (30)

\p{Script_Extensions: Mandaic} (Short: \p{Scx=Mand}, \p{Mand})
        (30: U+0640, U+0840..085B, U+085E)

\p{Script_Extensions: Mani} \p{Script_Extensions=Manichaean} (52)

\p{Script_Extensions: Manichaean} (Short: \p{Scx=Mani}, \p{Mani})
        (52: U+0640, U+10AC0..10AE6,
        U+10AEB..10AF6)

\p{Script_Extensions: Marc} \p{Script_Extensions=Marchen} (68)

\p{Script_Extensions: Marchen} (Short: \p{Scx=Marc}, \p{Marc})
        (68: U+11C70..11C8F, U+11C92..11CA7,
        U+11CA9..11CB6)

\p{Script_Extensions: Masaram_Gondi} (Short: \p{Scx=Gonm},
        \p{Gonm}) (77: U+0964..0965,
        U+11D00..11D06, U+11D08..11D09,
        U+11D0B..11D36, U+11D3A, U+11D3C..11D3D
        ...)

\p{Script_Extensions: Medefaidrin} (Short: \p{Scx=Medf}, \p{Medf})
        (91: U+16E40..16E9A)

\p{Script_Extensions: Medf} \p{Script_Extensions=Medefaidrin} (91)
\p{Script_Extensions: Meetei_Mayek} (Short: \p{Scx=Mtei},
           \p{Mtei}) (79: U+AAE0..AAF6,
           U+ABC0..ABED, U+ABF0..ABF9)
\p{Script_Extensions: Mend} \p{Script_Extensions=Mende_Kikakui}
           (213)
\p{Script_Extensions: Mende_Kikakui} (Short: \p{Scx=Mend},
           \p{Mend}) (213: U+1E800..1E8C4,
           U+1E8C7..1E8D6)
\p{Script_Extensions: Merc} \p{Script_Extensions=Meroitic_Cursive}
           (90)
\p{Script_Extensions: Mero} \p{Script_Extensions=
           Meroitic_Hieroglyphs} (32)
\p{Script_Extensions: Meroitic_Cursive} (Short: \p{Scx=Merc},
           \p{Merc}) (90: U+109A0..109B7,
           U+109BC..109CF, U+109D2..109FF)
\p{Script_Extensions: Meroitic_Hieroglyphs} (Short: \p{Scx=Mero},
           \p{Mero}) (32: U+10980..1099F)
\p{Script_Extensions: Miao} (Short: \p{Scx=Miao}, \p{Miao}) (149:
           U+16F00..16F4A, U+16F4F..16F87,
           U+16F8F..16F9F)
\p{Script_Extensions: Mlym} \p{Script_Extensions=Malayalam} (126)
\p{Script_Extensions: Modi} (Short: \p{Scx=Modi}, \p{Modi}) (89:
           U+A830..A839, U+11600..11644,
           U+11650..11659)
\p{Script_Extensions: Mong} \p{Script_Extensions=Mongolian} (171)
\p{Script_Extensions: Mongolian} (Short: \p{Scx=Mong}, \p{Mong})
           (171: U+1800..180E, U+1810..1819,
           U+1820..1878, U+1880..18AA, U+202F,
           U+11660..1166C)
\p{Script_Extensions: Mro} (Short: \p{Scx=Mro}, \p{Mro}) (43:
           U+16A40..16A5E, U+16A60..16A69,
           U+16A6E..16A6F)
\p{Script_Extensions: Mroo} \p{Script_Extensions=Mro} (43)
\p{Script_Extensions: Mtei} \p{Script_Extensions=Meetei_Mayek} (79)
\p{Script_Extensions: Mult} \p{Script_Extensions=Multani} (48)
\p{Script_Extensions: Multani} (Short: \p{Scx=Mult}, \p{Mult})
           (48: U+0A66..0A6F, U+11280..11286,
           U+11288, U+1128A..1128D, U+1128F..1129D,
           U+1129F..112A9)

\p{Script_Extensions: Myanmar} (Short: \p{Scx=Mymr}, \p{Mymr})
    (224: U+1000..109F, U+A92E,
    U+A9E0..A9FE, U+AA60..AA7F)
\p{Script_Extensions: Mymr} \p{Script_Extensions=Myanmar} (224)
\p{Script_Extensions: Nabataean} (Short: \p{Scx=Nbat}, \p{Nbat})
    (40: U+10880..1089E, U+108A7..108AF)
\p{Script_Extensions: Nand} \p{Script_Extensions=Nandinagari} (86)
\p{Script_Extensions: Nandinagari} (Short: \p{Scx=Nand}, \p{Nand})
    (86: U+0964..0965, U+0CE6..0CEF, U+1CE9,
    U+1CF2, U+1CFA, U+A830..A835 ...)
\p{Script_Extensions: Narb} \p{Script_Extensions=
    Old_North_Arabian} (32)
\p{Script_Extensions: Nbat} \p{Script_Extensions=Nabataean} (40)
\p{Script_Extensions: New_Tai_Lue} (Short: \p{Scx=Talu}, \p{Talu})
    (83: U+1980..19AB, U+19B0..19C9,
    U+19D0..19DA, U+19DE..19DF)
\p{Script_Extensions: Newa} (Short: \p{Scx=Newa}, \p{Newa}) (97:
    U+11400..1145B, U+1145D..11461)
\p{Script_Extensions: Nko} (Short: \p{Scx=Nko}, \p{Nko}) (62:
    U+07C0..07FA, U+07FD..07FF)
\p{Script_Extensions: Nkoo} \p{Script_Extensions=Nko} (62)
\p{Script_Extensions: Nshu} \p{Script_Extensions=Nushu} (397)
\p{Script_Extensions: Nushu} (Short: \p{Scx=Nshu}, \p{Nshu}) (397:
    U+16FE1, U+1B170..1B2FB)
\p{Script_Extensions: Nyiakeng_Puachue_Hmong} (Short: \p{Scx=
    Hmnp}, \p{Hmnp}) (71: U+1E100..1E12C,
    U+1E130..1E13D, U+1E140..1E149,
    U+1E14E..1E14F)
\p{Script_Extensions: Ogam} \p{Script_Extensions=Ogham} (29)
\p{Script_Extensions: Ogham} (Short: \p{Scx=Ogam}, \p{Ogam}) (29:
    U+1680..169C)
\p{Script_Extensions: Ol_Chiki} (Short: \p{Scx=Olck}, \p{Olck})
    (48: U+1C50..1C7F)
\p{Script_Extensions: Olck} \p{Script_Extensions=Ol_Chiki} (48)
\p{Script_Extensions: Old_Hungarian} (Short: \p{Scx=Hung},
    \p{Hung}) (108: U+10C80..10CB2,
    U+10CC0..10CF2, U+10CFA..10CFF)
\p{Script_Extensions: Old_Italic} (Short: \p{Scx=Ital}, \p{Ital})
    (39: U+10300..10323, U+1032D..1032F)
\p{Script_Extensions: Old_North_Arabian} (Short: \p{Scx=Narb},

\p{Narb}) (32: U+10A80..10A9F)
\p{Script_Extensions: Old_Permic} (Short: \p{Scx=Perm}, \p{Perm})
                (44: U+0483, U+10350..1037A)
\p{Script_Extensions: Old_Persian} (Short: \p{Scx=Xpeo}, \p{Xpeo})
                (50: U+103A0..103C3, U+103C8..103D5)
\p{Script_Extensions: Old_Sogdian} (Short: \p{Scx=Sogo}, \p{Sogo})
                (40: U+10F00..10F27)
\p{Script_Extensions: Old_South_Arabian} (Short: \p{Scx=Sarb},
                \p{Sarb}) (32: U+10A60..10A7F)
\p{Script_Extensions: Old_Turkic} (Short: \p{Scx=Orkh}, \p{Orkh})
                (73: U+10C00..10C48)
\p{Script_Extensions: Oriya} (Short: \p{Scx=Orya}, \p{Orya}) (97:
                U+0951..0952, U+0964..0965,
                U+0B01..0B03, U+0B05..0B0C,
                U+0B0F..0B10, U+0B13..0B28 ...)
\p{Script_Extensions: Orkh} \p{Script_Extensions=Old_Turkic} (73)
\p{Script_Extensions: Orya} \p{Script_Extensions=Oriya} (97)
\p{Script_Extensions: Osage} (Short: \p{Scx=Osge}, \p{Osge}) (72:
                U+104B0..104D3, U+104D8..104FB)
\p{Script_Extensions: Osge} \p{Script_Extensions=Osage} (72)
\p{Script_Extensions: Osma} \p{Script_Extensions=Osmanya} (40)
\p{Script_Extensions: Osmanya} (Short: \p{Scx=Osma}, \p{Osma})
                (40: U+10480..1049D, U+104A0..104A9)
\p{Script_Extensions: Pahawh_Hmong} (Short: \p{Scx=Hmng},
                \p{Hmng}) (127: U+16B00..16B45,
                U+16B50..16B59, U+16B5B..16B61,
                U+16B63..16B77, U+16B7D..16B8F)
\p{Script_Extensions: Palm} \p{Script_Extensions=Palmyrene} (32)
\p{Script_Extensions: Palmyrene} (Short: \p{Scx=Palm}, \p{Palm})
                (32: U+10860..1087F)
\p{Script_Extensions: Pau_Cin_Hau} (Short: \p{Scx=Pauc}, \p{Pauc})
                (57: U+11AC0..11AF8)
\p{Script_Extensions: Pauc} \p{Script_Extensions=Pau_Cin_Hau} (57)
\p{Script_Extensions: Perm} \p{Script_Extensions=Old_Permic} (44)
\p{Script_Extensions: Phag} \p{Script_Extensions=Phags_Pa} (59)
\p{Script_Extensions: Phags_Pa} (Short: \p{Scx=Phag}, \p{Phag})
                (59: U+1802..1803, U+1805, U+A840..A877)
\p{Script_Extensions: Phli} \p{Script_Extensions=
                Inscriptional_Pahlavi} (27)
\p{Script_Extensions: Phlp} \p{Script_Extensions=Psalter_Pahlavi}

(30)

\p{Script_Extensions: Phnx} \p{Script_Extensions=Phoenician} (29)

\p{Script_Extensions: Phoenician} (Short: \p{Scx=Phnx}, \p{Phnx})
        (29: U+10900..1091B, U+1091F)

\p{Script_Extensions: Plrd} \p{Script_Extensions=Miao} (149)

\p{Script_Extensions: Prti} \p{Script_Extensions=
        Inscriptional_Parthian} (30)

\p{Script_Extensions: Psalter_Pahlavi} (Short: \p{Scx=Phlp},
        \p{Phlp}) (30: U+0640, U+10B80..10B91,
        U+10B99..10B9C, U+10BA9..10BAF)

\p{Script_Extensions: Qaac} \p{Script_Extensions=Coptic} (165)

\p{Script_Extensions: Qaai} \p{Script_Extensions=Inherited} (503)

\p{Script_Extensions: Rejang} (Short: \p{Scx=Rjng}, \p{Rjng}) (37:
        U+A930..A953, U+A95F)

\p{Script_Extensions: Rjng} \p{Script_Extensions=Rejang} (37)

\p{Script_Extensions: Rohg} \p{Script_Extensions=Hanifi_Rohingya}
        (55)

\p{Script_Extensions: Runic} (Short: \p{Scx=Runr}, \p{Runr}) (86:
        U+16A0..16EA, U+16EE..16F8)

\p{Script_Extensions: Runr} \p{Script_Extensions=Runic} (86)

\p{Script_Extensions: Samaritan} (Short: \p{Scx=Samr}, \p{Samr})
        (61: U+0800..082D, U+0830..083E)

\p{Script_Extensions: Samr} \p{Script_Extensions=Samaritan} (61)

\p{Script_Extensions: Sarb} \p{Script_Extensions=
        Old_South_Arabian} (32)

\p{Script_Extensions: Saur} \p{Script_Extensions=Saurashtra} (82)

\p{Script_Extensions: Saurashtra} (Short: \p{Scx=Saur}, \p{Saur})
        (82: U+A880..A8C5, U+A8CE..A8D9)

\p{Script_Extensions: Sgnw} \p{Script_Extensions=SignWriting} (672)

\p{Script_Extensions: Sharada} (Short: \p{Scx=Shrd}, \p{Shrd})
        (102: U+0951, U+1CD7, U+1CD9,
        U+1CDC..1CDD, U+1CE0, U+11180..111DF)

\p{Script_Extensions: Shavian} (Short: \p{Scx=Shaw}, \p{Shaw})
        (48: U+10450..1047F)

\p{Script_Extensions: Shaw} \p{Script_Extensions=Shavian} (48)

\p{Script_Extensions: Shrd} \p{Script_Extensions=Sharada} (102)

\p{Script_Extensions: Sidd} \p{Script_Extensions=Siddham} (92)

\p{Script_Extensions: Siddham} (Short: \p{Scx=Sidd}, \p{Sidd})
        (92: U+11580..115B5, U+115B8..115DD)

\p{Script_Extensions: SignWriting} (Short: \p{Scx=Sgnw}, \p{Sgnw})

```
                        (672: U+1D800..1DA8B, U+1DA9B..1DA9F,
                        U+1DAA1..1DAAF)
\p{Script_Extensions: Sind} \p{Script_Extensions=Khudawadi} (81)
\p{Script_Extensions: Sinh} \p{Script_Extensions=Sinhala} (113)
\p{Script_Extensions: Sinhala} (Short: \p{Scx=Sinh}, \p{Sinh})
                        (113: U+0964..0965, U+0D81..0D83,
                        U+0D85..0D96, U+0D9A..0DB1,
                        U+0DB3..0DBB, U+0DBD ...)
\p{Script_Extensions: Sogd} \p{Script_Extensions=Sogdian} (43)
\p{Script_Extensions: Sogdian} (Short: \p{Scx=Sogd}, \p{Sogd})
                        (43: U+0640, U+10F30..10F59)
\p{Script_Extensions: Sogo} \p{Script_Extensions=Old_Sogdian} (40)
\p{Script_Extensions: Sora} \p{Script_Extensions=Sora_Sompeng} (35)
\p{Script_Extensions: Sora_Sompeng} (Short: \p{Scx=Sora},
                        \p{Sora}) (35: U+110D0..110E8,
                        U+110F0..110F9)
\p{Script_Extensions: Soyo} \p{Script_Extensions=Soyombo} (83)
\p{Script_Extensions: Soyombo} (Short: \p{Scx=Soyo}, \p{Soyo})
                        (83: U+11A50..11AA2)
\p{Script_Extensions: Sund} \p{Script_Extensions=Sundanese} (72)
\p{Script_Extensions: Sundanese} (Short: \p{Scx=Sund}, \p{Sund})
                        (72: U+1B80..1BBF, U+1CC0..1CC7)
\p{Script_Extensions: Sylo} \p{Script_Extensions=Syloti_Nagri} (57)
\p{Script_Extensions: Syloti_Nagri} (Short: \p{Scx=Sylo},
                        \p{Sylo}) (57: U+0964..0965,
                        U+09E6..09EF, U+A800..A82C)
\p{Script_Extensions: Syrc} \p{Script_Extensions=Syriac} (106)
\p{Script_Extensions: Syriac} (Short: \p{Scx=Syrc}, \p{Syrc})
                        (106: U+060C, U+061B..061C, U+061F,
                        U+0640, U+064B..0655, U+0670 ...)
\p{Script_Extensions: Tagalog} (Short: \p{Scx=Tglg}, \p{Tglg})
                        (22: U+1700..170C, U+170E..1714,
                        U+1735..1736)
\p{Script_Extensions: Tagb} \p{Script_Extensions=Tagbanwa} (20)
\p{Script_Extensions: Tagbanwa} (Short: \p{Scx=Tagb}, \p{Tagb})
                        (20: U+1735..1736, U+1760..176C,
                        U+176E..1770, U+1772..1773)
\p{Script_Extensions: Tai_Le} (Short: \p{Scx=Tale}, \p{Tale}) (45:
                        U+1040..1049, U+1950..196D, U+1970..1974)
\p{Script_Extensions: Tai_Tham} (Short: \p{Scx=Lana}, \p{Lana})
```


                (127: U+1A20..1A5E, U+1A60..1A7C,

                U+1A7F..1A89, U+1A90..1A99, U+1AA0..1AAD)

\p{Script_Extensions: Tai_Viet} (Short: \p{Scx=Tavt}, \p{Tavt})

                (72: U+AA80..AAC2, U+AADB..AADF)

\p{Script_Extensions: Takr} \p{Script_Extensions=Takri} (79)

\p{Script_Extensions: Takri} (Short: \p{Scx=Takr}, \p{Takr}) (79:

                U+0964..0965, U+A830..A839,

                U+11680..116B8, U+116C0..116C9)

\p{Script_Extensions: Tale} \p{Script_Extensions=Tai_Le} (45)

\p{Script_Extensions: Talu} \p{Script_Extensions=New_Tai_Lue} (83)

\p{Script_Extensions: Tamil} (Short: \p{Scx=Taml}, \p{Taml}) (133:

                U+0951..0952, U+0964..0965,

                U+0B82..0B83, U+0B85..0B8A,

                U+0B8E..0B90, U+0B92..0B95 ...)

\p{Script_Extensions: Taml} \p{Script_Extensions=Tamil} (133)

\p{Script_Extensions: Tang} \p{Script_Extensions=Tangut} (6914)

\p{Script_Extensions: Tangut} (Short: \p{Scx=Tang}, \p{Tang})

                (6914: U+16FE0, U+17000..187F7,

                U+18800..18AFF, U+18D00..18D08)

\p{Script_Extensions: Tavt} \p{Script_Extensions=Tai_Viet} (72)

\p{Script_Extensions: Telu} \p{Script_Extensions=Telugu} (104)

\p{Script_Extensions: Telugu} (Short: \p{Scx=Telu}, \p{Telu})

                (104: U+0951..0952, U+0964..0965,

                U+0C00..0C0C, U+0C0E..0C10,

                U+0C12..0C28, U+0C2A..0C39 ...)

\p{Script_Extensions: Tfng} \p{Script_Extensions=Tifinagh} (59)

\p{Script_Extensions: Tglg} \p{Script_Extensions=Tagalog} (22)

\p{Script_Extensions: Thaa} \p{Script_Extensions=Thaana} (66)

\p{Script_Extensions: Thaana} (Short: \p{Scx=Thaa}, \p{Thaa}) (66:

                U+060C, U+061B..061C, U+061F,

                U+0660..0669, U+0780..07B1, U+FDF2 ...)

\p{Script_Extensions: Thai} (Short: \p{Scx=Thai}, \p{Thai}) (86:

                U+0E01..0E3A, U+0E40..0E5B)

\p{Script_Extensions: Tibetan} (Short: \p{Scx=Tibt}, \p{Tibt})

                (207: U+0F00..0F47, U+0F49..0F6C,

                U+0F71..0F97, U+0F99..0FBC,

                U+0FBE..0FCC, U+0FCE..0FD4 ...)

\p{Script_Extensions: Tibt} \p{Script_Extensions=Tibetan} (207)

\p{Script_Extensions: Tifinagh} (Short: \p{Scx=Tfng}, \p{Tfng})

                (59: U+2D30..2D67, U+2D6F..2D70, U+2D7F)

\p{Script_Extensions: Tirh} \p{Script_Extensions=Tirhuta} (97)
\p{Script_Extensions: Tirhuta} (Short: \p{Scx=Tirh}, \p{Tirh})
                  (97: U+0951..0952, U+0964..0965, U+1CF2,
                  U+A830..A839, U+11480..114C7,
                  U+114D0..114D9)
\p{Script_Extensions: Ugar} \p{Script_Extensions=Ugaritic} (31)
\p{Script_Extensions: Ugaritic} (Short: \p{Scx=Ugar}, \p{Ugar})
                  (31: U+10380..1039D, U+1039F)
\p{Script_Extensions: Unknown} (Short: \p{Scx=Zzzz}, \p{Zzzz})
                  (970_188 plus all above-Unicode code
                  points: U+0378..0379, U+0380..0383,
                  U+038B, U+038D, U+03A2, U+0530 ...)
\p{Script_Extensions: Vai} (Short: \p{Scx=Vai}, \p{Vai}) (300:
                  U+A500..A62B)
\p{Script_Extensions: Vaii} \p{Script_Extensions=Vai} (300)
\p{Script_Extensions: Wancho} (Short: \p{Scx=Wcho}, \p{Wcho}) (59:
                  U+1E2C0..1E2F9, U+1E2FF)
\p{Script_Extensions: Wara} \p{Script_Extensions=Warang_Citi} (84)
\p{Script_Extensions: Warang_Citi} (Short: \p{Scx=Wara}, \p{Wara})
                  (84: U+118A0..118F2, U+118FF)
\p{Script_Extensions: Wcho} \p{Script_Extensions=Wancho} (59)
\p{Script_Extensions: Xpeo} \p{Script_Extensions=Old_Persian} (50)
\p{Script_Extensions: Xsux} \p{Script_Extensions=Cuneiform} (1234)
\p{Script_Extensions: Yezi} \p{Script_Extensions=Yezidi} (60)
\p{Script_Extensions: Yezidi} (Short: \p{Scx=Yezi}, \p{Yezi}) (60:
                  U+060C, U+061B, U+061F, U+0660..0669,
                  U+10E80..10EA9, U+10EAB..10EAD ...)
\p{Script_Extensions: Yi} (Short: \p{Scx=Yi}, \p{Yi}) (1246:
                  U+3001..3002, U+3008..3011,
                  U+3014..301B, U+30FB, U+A000..A48C,
                  U+A490..A4C6 ...)
\p{Script_Extensions: Yiii} \p{Script_Extensions=Yi} (1246)
\p{Script_Extensions: Zanabazar_Square} (Short: \p{Scx=Zanb},
                  \p{Zanb}) (72: U+11A00..11A47)
\p{Script_Extensions: Zanb} \p{Script_Extensions=Zanabazar_Square}
                  (72)
\p{Script_Extensions: Zinh} \p{Script_Extensions=Inherited} (503)
\p{Script_Extensions: Zyyy} \p{Script_Extensions=Common} (7661)
\p{Script_Extensions: Zzzz} \p{Script_Extensions=Unknown} (970_188
                  plus all above-Unicode code points)

\p{Scx: *}   \p{Script_Extensions: *}
\p{SD}   \p{Soft_Dotted} (= \p{Soft_Dotted=Y}) (46)
\p{SD: *}   \p{Soft_Dotted: *}
\p{Sentence_Break: AT} \p{Sentence_Break=ATerm} (4)
\p{Sentence_Break: ATerm} (Short: \p{SB=AT}) (4: [.], U+2024,
                U+FE52, U+FF0E)
\p{Sentence_Break: CL} \p{Sentence_Break=Close} (187)
\p{Sentence_Break: Close} (Short: \p{SB=CL}) (187: [\"\'\(\)\[\]
                \{\}\xab\xbb], U+0F3A..0F3D,
                U+169B..169C, U+2018..201F,
                U+2039..203A, U+2045..2046 ...)
\p{Sentence_Break: CR} (Short: \p{SB=CR}) (1: [\r])
\p{Sentence_Break: EX} \p{Sentence_Break=Extend} (2395)
\p{Sentence_Break: Extend} (Short: \p{SB=EX}) (2395: U+0300..036F,
                U+0483..0489, U+0591..05BD, U+05BF,
                U+05C1..05C2, U+05C4..05C5 ...)
\p{Sentence_Break: FO} \p{Sentence_Break=Format} (63)
\p{Sentence_Break: Format} (Short: \p{SB=FO}) (63: [\xad],
                U+0600..0605, U+061C, U+06DD, U+070F,
                U+08E2 ...)
\p{Sentence_Break: LE} \p{Sentence_Break=OLetter} (127_413)
\p{Sentence_Break: LF} (Short: \p{SB=LF}) (1: [\n])
\p{Sentence_Break: LO} \p{Sentence_Break=Lower} (2297)
\p{Sentence_Break: Lower} (Short: \p{SB=LO}) (2297: [a-z\xaa\xb5
                \xba\xdf-\xf6\xf8-\xff], U+0101, U+0103,
                U+0105, U+0107, U+0109 ...)
\p{Sentence_Break: NU} \p{Sentence_Break=Numeric} (652)
\p{Sentence_Break: Numeric} (Short: \p{SB=NU}) (652: [0-9],
                U+0660..0669, U+066B..066C,
                U+06F0..06F9, U+07C0..07C9, U+0966..096F
                ...)
\p{Sentence_Break: OLetter} (Short: \p{SB=LE}) (127_413: U+01BB,
                U+01C0..01C3, U+0294, U+02B9..02BF,
                U+02C6..02D1, U+02EC ...)
\p{Sentence_Break: Other} (Short: \p{SB=XX}) (979_014 plus all
                above-Unicode code points: [^\t\n\cK\f
                \r\x20!\"\'\(\),\-.0-9:?A-Z\[\]a-z\{\}
                \x85\xa0\xaa-\xab\xad\xb5\xba-\xbb\xc0-
                \xd6\xd8-\xf6\xf8-\xff], U+02C2..02C5,
                U+02D2..02DF, U+02E5..02EB, U+02ED,

U+02EF..02FF ...)

\p{Sentence_Break: SC} \p{Sentence_Break=SContinue} (26)

\p{Sentence_Break: SContinue} (Short: \p{SB=SC}) (26: [,\-:],
                U+055D, U+060C..060D, U+07F8, U+1802,
                U+1808 ...)

\p{Sentence_Break: SE} \p{Sentence_Break=Sep} (3)

\p{Sentence_Break: Sep} (Short: \p{SB=SE}) (3: [\x85],
                U+2028..2029)

\p{Sentence_Break: Sp} (Short: \p{SB=Sp}) (20: [\t\cK\f\x20\xa0],
                U+1680, U+2000..200A, U+202F, U+205F,
                U+3000)

\p{Sentence_Break: ST} \p{Sentence_Break=STerm} (140)

\p{Sentence_Break: STerm} (Short: \p{SB=ST}) (140: [!?], U+0589,
                U+061E..061F, U+06D4, U+0700..0702,
                U+07F9 ...)

\p{Sentence_Break: UP} \p{Sentence_Break=Upper} (1896)

\p{Sentence_Break: Upper} (Short: \p{SB=UP}) (1896: [A-Z\xc0-\xd6
                \xd8-\xde], U+0100, U+0102, U+0104,
                U+0106, U+0108 ...)

\p{Sentence_Break: XX} \p{Sentence_Break=Other} (979_014 plus all
                above-Unicode code points)

\p{Sentence_Terminal}   \p{Sentence_Terminal=Y} (Short: \p{STerm})
                (143)

\p{Sentence_Terminal: N*} (Short: \p{STerm=N}, \P{STerm})
                (1_113_969 plus all above-Unicode code
                points: [\x00-\x20\"#\$\%&\'\(\)*+,\-
                \/0-9:;<=>\@A-Z\[\\\]\^_'a-z\{\|\}~\x7f-
                \xff], U+0100..0588, U+058A..061D,
                U+0620..06D3, U+06D5..06FF, U+0703..07F8
                ...)

\p{Sentence_Terminal: Y*} (Short: \p{STerm=Y}, \p{STerm}) (143:
                [!.?], U+0589, U+061E..061F, U+06D4,
                U+0700..0702, U+07F9 ...)

\p{Separator}   \p{General_Category=Separator} (Short:
                \p{Z}) (19)

\p{Sgnw}   \p{SignWriting} (= \p{Script_Extensions=
                SignWriting}) (672)

\p{Sharada}   \p{Script_Extensions=Sharada} (Short:
                \p{Shrd}; NOT \p{Block=Sharada}) (102)

\p{Shavian}   \p{Script_Extensions=Shavian} (Short:

```
                     \p{Shaw}) (48)
  \p{Shaw}              \p{Shavian} (= \p{Script_Extensions=
                     Shavian}) (48)
 X \p{Shorthand_Format_Controls} \p{Block=Shorthand_Format_Controls}
                     (16)
  \p{Shrd}              \p{Sharada} (= \p{Script_Extensions=
                     Sharada}) (NOT \p{Block=Sharada}) (102)
  \p{Sidd}              \p{Siddham} (= \p{Script_Extensions=
                     Siddham}) (NOT \p{Block=Siddham}) (92)
  \p{Siddham}            \p{Script_Extensions=Siddham} (Short:
                     \p{Sidd}; NOT \p{Block=Siddham}) (92)
  \p{SignWriting}        \p{Script_Extensions=SignWriting} (Short:
                     \p{Sgnw}) (672)
  \p{Sind}              \p{Khudawadi} (= \p{Script_Extensions=
                     Khudawadi}) (NOT \p{Block=Khudawadi})
                     (81)
  \p{Sinh}              \p{Sinhala} (= \p{Script_Extensions=
                     Sinhala}) (NOT \p{Block=Sinhala}) (113)
  \p{Sinhala}            \p{Script_Extensions=Sinhala} (Short:
                     \p{Sinh}; NOT \p{Block=Sinhala}) (113)
 X \p{Sinhala_Archaic_Numbers} \p{Block=Sinhala_Archaic_Numbers} (32)
  \p{Sk}               \p{Modifier_Symbol} (=
                     \p{General_Category=Modifier_Symbol})
                     (123)
  \p{Sm}               \p{Math_Symbol} (= \p{General_Category=
                     Math_Symbol}) (948)
 X \p{Small_Form_Variants} \p{Block=Small_Form_Variants} (Short:
                     \p{InSmallForms}) (32)
 X \p{Small_Forms}         \p{Small_Form_Variants} (= \p{Block=
                     Small_Form_Variants}) (32)
 X \p{Small_Kana_Ext}      \p{Small_Kana_Extension} (= \p{Block=
                     Small_Kana_Extension}) (64)
 X \p{Small_Kana_Extension} \p{Block=Small_Kana_Extension} (Short:
                     \p{InSmallKanaExt}) (64)
  \p{So}               \p{Other_Symbol} (= \p{General_Category=
                     Other_Symbol}) (6431)
  \p{Soft_Dotted}        \p{Soft_Dotted=Y} (Short: \p{SD}) (46)
  \p{Soft_Dotted: N*}     (Short: \p{SD=N}, \P{SD}) (1_114_066 plus
                     all above-Unicode code points: [\x00-
                     \x20!\"#\$\%&\'\(\)*+,\-.\/0-9:;<=>?\@A-
```

```
                  Z\[\\\]\^_'a-hk-z\{\|\}~\x7f-\xff],
                  U+0100..012E, U+0130..0248,
                  U+024A..0267, U+0269..029C, U+029E..02B1
                  ...)
```

\p{Soft_Dotted: Y*}   (Short: \p{SD=Y}, \p{SD}) (46: [i-j],
                  U+012F, U+0249, U+0268, U+029D, U+02B2
                  ...)

\p{Sogd}   \p{Sogdian} (= \p{Script_Extensions=
                  Sogdian}) (NOT \p{Block=Sogdian}) (43)

\p{Sogdian}   \p{Script_Extensions=Sogdian} (Short:
                  \p{Sogd}; NOT \p{Block=Sogdian}) (43)

\p{Sogo}   \p{Old_Sogdian} (= \p{Script_Extensions=
                  Old_Sogdian}) (NOT \p{Block=
                  Old_Sogdian}) (40)

\p{Sora}   \p{Sora_Sompeng} (= \p{Script_Extensions=
                  Sora_Sompeng}) (NOT \p{Block=
                  Sora_Sompeng}) (35)

\p{Sora_Sompeng}   \p{Script_Extensions=Sora_Sompeng} (Short:
                  \p{Sora}; NOT \p{Block=Sora_Sompeng})
                  (35)

\p{Soyo}   \p{Soyombo} (= \p{Script_Extensions=
                  Soyombo}) (NOT \p{Block=Soyombo}) (83)

\p{Soyombo}   \p{Script_Extensions=Soyombo} (Short:
                  \p{Soyo}; NOT \p{Block=Soyombo}) (83)

\p{Space}   \p{White_Space} (= \p{White_Space=Y}) (25)

\p{Space: *}   \p{White_Space: *}

\p{Space_Separator}   \p{General_Category=Space_Separator}
                  (Short: \p{Zs}) (17)

\p{SpacePerl}   \p{XPosixSpace} (25)

\p{Spacing_Mark}   \p{General_Category=Spacing_Mark} (Short:
                  \p{Mc}) (443)

X \p{Spacing_Modifier_Letters} \p{Block=Spacing_Modifier_Letters}
                  (Short: \p{InModifierLetters}) (80)

X \p{Specials}   \p{Block=Specials} (16)

\p{STerm}   \p{Sentence_Terminal} (=
                  \p{Sentence_Terminal=Y}) (143)

\p{STerm: *}   \p{Sentence_Terminal: *}

\p{Sund}   \p{Sundanese} (= \p{Script_Extensions=
                  Sundanese}) (NOT \p{Block=Sundanese})
                  (72)

  \p{Sundanese}   \p{Script_Extensions=Sundanese} (Short:
                \p{Sund}; NOT \p{Block=Sundanese}) (72)

X \p{Sundanese_Sup}   \p{Sundanese_Supplement} (= \p{Block=
                Sundanese_Supplement}) (16)

X \p{Sundanese_Supplement} \p{Block=Sundanese_Supplement} (Short:
                \p{InSundaneseSup}) (16)

X \p{Sup_Arrows_A}   \p{Supplemental_Arrows_A} (= \p{Block=
                Supplemental_Arrows_A}) (16)

X \p{Sup_Arrows_B}   \p{Supplemental_Arrows_B} (= \p{Block=
                Supplemental_Arrows_B}) (128)

X \p{Sup_Arrows_C}   \p{Supplemental_Arrows_C} (= \p{Block=
                Supplemental_Arrows_C}) (256)

X \p{Sup_Math_Operators} \p{Supplemental_Mathematical_Operators} (=
                \p{Block=
                Supplemental_Mathematical_Operators})
                (256)

X \p{Sup_PUA_A}   \p{Supplementary_Private_Use_Area_A} (=
                \p{Block=
                Supplementary_Private_Use_Area_A})
                (65_536)

X \p{Sup_PUA_B}   \p{Supplementary_Private_Use_Area_B} (=
                \p{Block=
                Supplementary_Private_Use_Area_B})
                (65_536)

X \p{Sup_Punctuation}   \p{Supplemental_Punctuation} (= \p{Block=
                Supplemental_Punctuation}) (128)

X \p{Sup_Symbols_And_Pictographs}
                \p{Supplemental_Symbols_And_Pictographs}
                (= \p{Block=
                Supplemental_Symbols_And_Pictographs})
                (256)

X \p{Super_And_Sub}   \p{Superscripts_And_Subscripts} (=
                \p{Block=Superscripts_And_Subscripts})
                (48)

X \p{Superscripts_And_Subscripts} \p{Block=
                Superscripts_And_Subscripts} (Short:
                \p{InSuperAndSub}) (48)

X \p{Supplemental_Arrows_A} \p{Block=Supplemental_Arrows_A} (Short:
                \p{InSupArrowsA}) (16)

X \p{Supplemental_Arrows_B} \p{Block=Supplemental_Arrows_B} (Short:

```
                    \p{InSupArrowsB}) (128)
    X \p{Supplemental_Arrows_C} \p{Block=Supplemental_Arrows_C} (Short:
                    \p{InSupArrowsC}) (256)
    X \p{Supplemental_Mathematical_Operators} \p{Block=
                    Supplemental_Mathematical_Operators}
                    (Short: \p{InSupMathOperators}) (256)
    X \p{Supplemental_Punctuation} \p{Block=Supplemental_Punctuation}
                    (Short: \p{InSupPunctuation}) (128)
    X \p{Supplemental_Symbols_And_Pictographs} \p{Block=
                    Supplemental_Symbols_And_Pictographs}
                    (Short: \p{InSupSymbolsAndPictographs})
                    (256)
    X \p{Supplementary_Private_Use_Area_A} \p{Block=
                    Supplementary_Private_Use_Area_A}
                    (Short: \p{InSupPUAA}) (65_536)
    X \p{Supplementary_Private_Use_Area_B} \p{Block=
                    Supplementary_Private_Use_Area_B}
                    (Short: \p{InSupPUAB}) (65_536)
     \p{Surrogate}         \p{General_Category=Surrogate} (Short:
                    \p{Cs}) (2048)
    X \p{Sutton_SignWriting}  \p{Block=Sutton_SignWriting} (688)
     \p{Sylo}              \p{Syloti_Nagri} (= \p{Script_Extensions=
                    Syloti_Nagri}) (NOT \p{Block=
                    Syloti_Nagri}) (57)
     \p{Syloti_Nagri}      \p{Script_Extensions=Syloti_Nagri} (Short:
                    \p{Sylo}; NOT \p{Block=Syloti_Nagri})
                    (57)
     \p{Symbol}           \p{General_Category=Symbol} (Short: \p{S})
                    (7564)
    X \p{Symbols_And_Pictographs_Ext_A}
                    \p{Symbols_And_Pictographs_Extended_A}
                    (= \p{Block=
                    Symbols_And_Pictographs_Extended_A})
                    (144)
    X \p{Symbols_And_Pictographs_Extended_A} \p{Block=
                    Symbols_And_Pictographs_Extended_A} (144)
    X \p{Symbols_For_Legacy_Computing} \p{Block=
                    Symbols_For_Legacy_Computing} (256)
     \p{Syrc}             \p{Syriac} (= \p{Script_Extensions=
                    Syriac}) (NOT \p{Block=Syriac}) (106)
```

```
 \p{Syriac}           \p{Script_Extensions=Syriac} (Short:
                      \p{Syrc}; NOT \p{Block=Syriac}) (106)
X \p{Syriac_Sup}      \p{Syriac_Supplement} (= \p{Block=
                      Syriac_Supplement}) (16)
X \p{Syriac_Supplement}  \p{Block=Syriac_Supplement} (Short:
                      \p{InSyriacSup}) (16)
 \p{Tagalog}          \p{Script_Extensions=Tagalog} (Short:
                      \p{Tglg}; NOT \p{Block=Tagalog}) (22)
 \p{Tagb}             \p{Tagbanwa} (= \p{Script_Extensions=
                      Tagbanwa}) (NOT \p{Block=Tagbanwa}) (20)
 \p{Tagbanwa}         \p{Script_Extensions=Tagbanwa} (Short:
                      \p{Tagb}; NOT \p{Block=Tagbanwa}) (20)
X \p{Tags}            \p{Block=Tags} (128)
 \p{Tai_Le}           \p{Script_Extensions=Tai_Le} (Short:
                      \p{Tale}; NOT \p{Block=Tai_Le}) (45)
 \p{Tai_Tham}         \p{Script_Extensions=Tai_Tham} (Short:
                      \p{Lana}; NOT \p{Block=Tai_Tham}) (127)
 \p{Tai_Viet}         \p{Script_Extensions=Tai_Viet} (Short:
                      \p{Tavt}; NOT \p{Block=Tai_Viet}) (72)
X \p{Tai_Xuan_Jing}   \p{Tai_Xuan_Jing_Symbols} (= \p{Block=
                      Tai_Xuan_Jing_Symbols}) (96)
X \p{Tai_Xuan_Jing_Symbols} \p{Block=Tai_Xuan_Jing_Symbols} (Short:
                      \p{InTaiXuanJing}) (96)
 \p{Takr}             \p{Takri} (= \p{Script_Extensions=Takri})
                      (NOT \p{Block=Takri}) (79)
 \p{Takri}            \p{Script_Extensions=Takri} (Short:
                      \p{Takr}; NOT \p{Block=Takri}) (79)
 \p{Tale}             \p{Tai_Le} (= \p{Script_Extensions=
                      Tai_Le}) (NOT \p{Block=Tai_Le}) (45)
 \p{Talu}             \p{New_Tai_Lue} (= \p{Script_Extensions=
                      New_Tai_Lue}) (NOT \p{Block=
                      New_Tai_Lue}) (83)
 \p{Tamil}            \p{Script_Extensions=Tamil} (Short:
                      \p{Taml}; NOT \p{Block=Tamil}) (133)
X \p{Tamil_Sup}       \p{Tamil_Supplement} (= \p{Block=
                      Tamil_Supplement}) (64)
X \p{Tamil_Supplement}   \p{Block=Tamil_Supplement} (Short:
                      \p{InTamilSup}) (64)
 \p{Taml}             \p{Tamil} (= \p{Script_Extensions=Tamil})
                      (NOT \p{Block=Tamil}) (133)
```

```
 \p{Tang}           \p{Tangut} (= \p{Script_Extensions=
                    Tangut}) (NOT \p{Block=Tangut}) (6914)
 \p{Tangut}          \p{Script_Extensions=Tangut} (Short:
                    \p{Tang}; NOT \p{Block=Tangut}) (6914)
X \p{Tangut_Components}   \p{Block=Tangut_Components} (768)
X \p{Tangut_Sup}       \p{Tangut_Supplement} (= \p{Block=
                    Tangut_Supplement}) (144)
X \p{Tangut_Supplement}   \p{Block=Tangut_Supplement} (Short:
                    \p{InTangutSup}) (144)
 \p{Tavt}           \p{Tai_Viet} (= \p{Script_Extensions=
                    Tai_Viet}) (NOT \p{Block=Tai_Viet}) (72)
 \p{Telu}           \p{Telugu} (= \p{Script_Extensions=
                    Telugu}) (NOT \p{Block=Telugu}) (104)
 \p{Telugu}          \p{Script_Extensions=Telugu} (Short:
                    \p{Telu}; NOT \p{Block=Telugu}) (104)
 \p{Term}           \p{Terminal_Punctuation} (=
                    \p{Terminal_Punctuation=Y}) (267)
 \p{Term: *}         \p{Terminal_Punctuation: *}
 \p{Terminal_Punctuation} \p{Terminal_Punctuation=Y} (Short:
                    \p{Term}) (267)
 \p{Terminal_Punctuation: N*} (Short: \p{Term=N}, \P{Term})
                    (1_113_845 plus all above-Unicode code
                    points: [\x00-\x20\"#\$\%&\'\(\)*+\-\/0-
                    9<=>\@A-Z\[\\\]\^_'a-z\{\|\}~\x7f-\xff],
                    U+0100..037D, U+037F..0386,
                    U+0388..0588, U+058A..05C2, U+05C4..060B
                    ...)
 \p{Terminal_Punctuation: Y*} (Short: \p{Term=Y}, \p{Term}) (267:
                    [!,.:;?], U+037E, U+0387, U+0589,
                    U+05C3, U+060C ...)
 \p{Tfng}           \p{Tifinagh} (= \p{Script_Extensions=
                    Tifinagh}) (NOT \p{Block=Tifinagh}) (59)
 \p{Tglg}           \p{Tagalog} (= \p{Script_Extensions=
                    Tagalog}) (NOT \p{Block=Tagalog}) (22)
 \p{Thaa}           \p{Thaana} (= \p{Script_Extensions=
                    Thaana}) (NOT \p{Block=Thaana}) (66)
 \p{Thaana}          \p{Script_Extensions=Thaana} (Short:
                    \p{Thaa}; NOT \p{Block=Thaana}) (66)
 \p{Thai}           \p{Script_Extensions=Thai} (NOT \p{Block=
                    Thai}) (86)
```

\p{Tibetan}   \p{Script_Extensions=Tibetan} (Short:
                      \p{Tibt}; NOT \p{Block=Tibetan}) (207)
\p{Tibt}   \p{Tibetan} (= \p{Script_Extensions=
                      Tibetan}) (NOT \p{Block=Tibetan}) (207)
\p{Tifinagh}   \p{Script_Extensions=Tifinagh} (Short:
                      \p{Tfng}; NOT \p{Block=Tifinagh}) (59)
\p{Tirh}   \p{Tirhuta} (= \p{Script_Extensions=
                      Tirhuta}) (NOT \p{Block=Tirhuta}) (97)
\p{Tirhuta}   \p{Script_Extensions=Tirhuta} (Short:
                      \p{Tirh}; NOT \p{Block=Tirhuta}) (97)
\p{Title}   \p{Titlecase} (/i= Cased=Yes) (31)
\p{Titlecase}   (= \p{Gc=Lt}) (Short: \p{Title}; /i=
                      Cased=Yes) (31: U+01C5, U+01C8, U+01CB,
                      U+01F2, U+1F88..1F8F, U+1F98..1F9F ...)
\p{Titlecase_Letter}   \p{General_Category=Titlecase_Letter}
                      (Short: \p{Lt}; /i= General_Category=
                      Cased_Letter) (31)
X \p{Transport_And_Map}   \p{Transport_And_Map_Symbols} (= \p{Block=
                      Transport_And_Map_Symbols}) (128)
X \p{Transport_And_Map_Symbols} \p{Block=Transport_And_Map_Symbols}
                      (Short: \p{InTransportAndMap}) (128)
X \p{UCAS}   \p{Unified_Canadian_Aboriginal_Syllabics}
                      (= \p{Block=
                      Unified_Canadian_Aboriginal_Syllabics})
                      (640)
X \p{UCAS_Ext}   \p{Unified_Canadian_Aboriginal_Syllabics_-
                      Extended} (= \p{Block=
                      Unified_Canadian_Aboriginal_Syllabics_-
                      Extended}) (80)
\p{Ugar}   \p{Ugaritic} (= \p{Script_Extensions=
                      Ugaritic}) (NOT \p{Block=Ugaritic}) (31)
\p{Ugaritic}   \p{Script_Extensions=Ugaritic} (Short:
                      \p{Ugar}; NOT \p{Block=Ugaritic}) (31)
\p{UIdeo}   \p{Unified_Ideograph} (=
                      \p{Unified_Ideograph=Y}) (92_856)
\p{UIdeo: *}   \p{Unified_Ideograph: *}
\p{Unassigned}   \p{General_Category=Unassigned} (Short:
                      \p{Cn}) (830_672 plus all above-Unicode
                      code points)
\p{Unicode}   \p{Any} (1_114_112)

X \p{Unified_Canadian_Aboriginal_Syllabics} \p{Block=
           Unified_Canadian_Aboriginal_Syllabics}
           (Short: \p{InUCAS}) (640)
X \p{Unified_Canadian_Aboriginal_Syllabics_Extended} \p{Block=
           Unified_Canadian_Aboriginal_Syllabics_-
           Extended} (Short: \p{InUCASExt}) (80)
  \p{Unified_Ideograph}   \p{Unified_Ideograph=Y} (Short: \p{UIdeo})
           (92_856)
  \p{Unified_Ideograph: N*} (Short: \p{UIdeo=N}, \P{UIdeo})
           (1_021_256 plus all above-Unicode code
           points: U+0000..33FF, U+4DC0..4DFF,
           U+9FFD..FA0D, U+FA10, U+FA12,
           U+FA15..FA1E ...)
  \p{Unified_Ideograph: Y*} (Short: \p{UIdeo=Y}, \p{UIdeo}) (92_856:
           U+3400..4DBF, U+4E00..9FFC,
           U+FA0E..FA0F, U+FA11, U+FA13..FA14,
           U+FA1F ...)
  \p{Unknown}   \p{Script_Extensions=Unknown} (Short:
           \p{Zzzz}) (970_188 plus all above-
           Unicode code points)
  \p{Upper}   \p{XPosixUpper} (= \p{Uppercase=Y}) (/i=
           Cased=Yes) (1911)
  \p{Upper: *}   \p{Uppercase: *}
  \p{Uppercase}   \p{XPosixUpper} (= \p{Uppercase=Y}) (/i=
           Cased=Yes) (1911)
  \p{Uppercase: N*}   (Short: \p{Upper=N}, \P{Upper}; /i= Cased=
           No) (1_112_201 plus all above-Unicode
           code points: [\x00-\x20!\"#\$\%&\'
           \(\)*+,\-.\/0-9:;<=>?\@\[\\\]\^_'a-z\{
           \|\}~\x7f-\xbf\xd7\xdf-\xff], U+0101,
           U+0103, U+0105, U+0107, U+0109 ...)
  \p{Uppercase: Y*}   (Short: \p{Upper=Y}, \p{Upper}; /i= Cased=
           Yes) (1911: [A-Z\xc0-\xd6\xd8-\xde],
           U+0100, U+0102, U+0104, U+0106, U+0108
           ...)
  \p{Uppercase_Letter}   \p{General_Category=Uppercase_Letter}
           (Short: \p{Lu}; /i= General_Category=
           Cased_Letter) (1791)
  \p{Vai}   \p{Script_Extensions=Vai} (NOT \p{Block=
           Vai}) (300)

        \p{Vaii}   \p{Vai} (= \p{Script_Extensions=Vai}) (NOT
                              \p{Block=Vai}) (300)
    \p{Variation_Selector} \p{Variation_Selector=Y} (Short: \p{VS};
                              NOT \p{Variation_Selectors}) (259)
    \p{Variation_Selector: N*} (Short: \p{VS=N}, \P{VS}) (1_113_853
                              plus all above-Unicode code points:
                              U+0000..180A, U+180E..FDFF,
                              U+FE10..E00FF, U+E01F0..infinity)
    \p{Variation_Selector: Y*} (Short: \p{VS=Y}, \p{VS}) (259:
                              U+180B..180D, U+FE00..FE0F,
                              U+E0100..E01EF)
    X \p{Variation_Selectors} \p{Block=Variation_Selectors} (Short:
                              \p{InVS}) (16)
    X \p{Variation_Selectors_Supplement} \p{Block=
                              Variation_Selectors_Supplement} (Short:
                              \p{InVSSup}) (240)
    X \p{Vedic_Ext}   \p{Vedic_Extensions} (= \p{Block=
                              Vedic_Extensions}) (48)
    X \p{Vedic_Extensions}   \p{Block=Vedic_Extensions} (Short:
                              \p{InVedicExt}) (48)
    X \p{Vertical_Forms}   \p{Block=Vertical_Forms} (16)
     \p{Vertical_Orientation: R} \p{Vertical_Orientation=Rotated}
                              (786_865 plus all above-Unicode code
                              points)
     \p{Vertical_Orientation: Rotated} (Short: \p{Vo=R}) (786_865 plus
                              all above-Unicode code points: [\x00-
                              \xa6\xa8\xaa-\xad\xaf-\xb0\xb2-\xbb\xbf-
                              \xd6\xd8-\xf6\xf8-\xff], U+0100..02E9,
                              U+02EC..10FF, U+1200..1400,
                              U+1680..18AF, U+1900..2015 ...)
     \p{Vertical_Orientation: Tr} \p{Vertical_Orientation=
                              Transformed_Rotated} (47)
     \p{Vertical_Orientation: Transformed_Rotated} (Short: \p{Vo=Tr})
                              (47: U+2329..232A, U+3008..3011,
                              U+3014..301F, U+3030, U+30A0, U+30FC ...)
     \p{Vertical_Orientation: Transformed_Upright} (Short: \p{Vo=Tu})
                              (148: U+3001..3002, U+3041, U+3043,
                              U+3045, U+3047, U+3049 ...)
     \p{Vertical_Orientation: Tu} \p{Vertical_Orientation=
                              Transformed_Upright} (148)

\p{Vertical_Orientation: U} \p{Vertical_Orientation=Upright}
                (327_052)
\p{Vertical_Orientation: Upright} (Short: \p{Vo=U}) (327_052:
                [\xa7\xa9\xae\xb1\xbc-\xbe\xd7\xf7],
                U+02EA..02EB, U+1100..11FF,
                U+1401..167F, U+18B0..18FF, U+2016 ...)
\p{VertSpace}   \v (7: [\n\cK\f\r\x85], U+2028..2029)
\p{Vo: *}   \p{Vertical_Orientation: *}
\p{VS}   \p{Variation_Selector} (=
                \p{Variation_Selector=Y}) (NOT
                \p{Variation_Selectors}) (259)
\p{VS: *}   \p{Variation_Selector: *}
 X \p{VS_Sup}   \p{Variation_Selectors_Supplement} (=
                \p{Block=
                Variation_Selectors_Supplement}) (240)
\p{Wancho}   \p{Script_Extensions=Wancho} (Short:
                \p{Wcho}; NOT \p{Block=Wancho}) (59)
\p{Wara}   \p{Warang_Citi} (= \p{Script_Extensions=
                Warang_Citi}) (NOT \p{Block=
                Warang_Citi}) (84)
\p{Warang_Citi}   \p{Script_Extensions=Warang_Citi} (Short:
                \p{Wara}; NOT \p{Block=Warang_Citi}) (84)
\p{WB: *}   \p{Word_Break: *}
\p{Wcho}   \p{Wancho} (= \p{Script_Extensions=
                Wancho}) (NOT \p{Block=Wancho}) (59)
\p{White_Space}   \p{White_Space=Y} (Short: \p{Space}) (25)
\p{White_Space: N*}   (Short: \p{Space=N}, \P{Space}) (1_114_087
                plus all above-Unicode code points: [^
                \t\n\cK\f\r\x20\x85\xa0], U+0100..167F,
                U+1681..1FFF, U+200B..2027,
                U+202A..202E, U+2030..205E ...)
\p{White_Space: Y*}   (Short: \p{Space=Y}, \p{Space}) (25: [\t
                \n\cK\f\r\x20\x85\xa0], U+1680,
                U+2000..200A, U+2028..2029, U+202F,
                U+205F ...)
\p{Word}   \p{XPosixWord} (134_564)
\p{Word_Break: ALetter} (Short: \p{WB=LE}) (28_854: [A-Za-z\xaa
                \xb5\xba\xc0-\xd6\xd8-\xf6\xf8-\xff],
                U+0100..02D7, U+02DE..02FF,
                U+0370..0374, U+0376..0377, U+037A..037D

                       ...)
     \p{Word_Break: CR}   (Short: \p{WB=CR}) (1: [\r])
     \p{Word_Break: Double_Quote} (Short: \p{WB=DQ}) (1: [\"])
     \p{Word_Break: DQ}   \p{Word_Break=Double_Quote} (1)
     \p{Word_Break: E_Base} (Short: \p{WB=EB}) (0)
     \p{Word_Break: E_Base_GAZ} (Short: \p{WB=EBG}) (0)
     \p{Word_Break: E_Modifier} (Short: \p{WB=EM}) (0)
     \p{Word_Break: EB}   \p{Word_Break=E_Base} (0)
     \p{Word_Break: EBG}   \p{Word_Break=E_Base_GAZ} (0)
     \p{Word_Break: EM}   \p{Word_Break=E_Modifier} (0)
     \p{Word_Break: EX}   \p{Word_Break=ExtendNumLet} (11)
     \p{Word_Break: Extend} (Short: \p{WB=Extend}) (2399:
                       U+0300..036F, U+0483..0489,
                       U+0591..05BD, U+05BF, U+05C1..05C2,
                       U+05C4..05C5 ...)
     \p{Word_Break: ExtendNumLet} (Short: \p{WB=EX}) (11: [_], U+202F,
                       U+203F..2040, U+2054, U+FE33..FE34,
                       U+FE4D..FE4F ...)
     \p{Word_Break: FO}   \p{Word_Break=Format} (62)
     \p{Word_Break: Format} (Short: \p{WB=FO}) (62: [\xad],
                       U+0600..0605, U+061C, U+06DD, U+070F,
                       U+08E2 ...)
     \p{Word_Break: GAZ}   \p{Word_Break=Glue_After_Zwj} (0)
     \p{Word_Break: Glue_After_Zwj} (Short: \p{WB=GAZ}) (0)
     \p{Word_Break: Hebrew_Letter} (Short: \p{WB=HL}) (75:
                       U+05D0..05EA, U+05EF..05F2, U+FB1D,
                       U+FB1F..FB28, U+FB2A..FB36, U+FB38..FB3C
                       ...)
     \p{Word_Break: HL}   \p{Word_Break=Hebrew_Letter} (75)
     \p{Word_Break: KA}   \p{Word_Break=Katakana} (314)
     \p{Word_Break: Katakana} (Short: \p{WB=KA}) (314: U+3031..3035,
                       U+309B..309C, U+30A0..30FA,
                       U+30FC..30FF, U+31F0..31FF, U+32D0..32FE
                       ...)
     \p{Word_Break: LE}   \p{Word_Break=ALetter} (28_854)
     \p{Word_Break: LF}   (Short: \p{WB=LF}) (1: [\n])
     \p{Word_Break: MB}   \p{Word_Break=MidNumLet} (7)
     \p{Word_Break: MidLetter} (Short: \p{WB=ML}) (9: [:\xb7], U+0387,
                       U+055F, U+05F4, U+2027, U+FE13 ...)
     \p{Word_Break: MidNum} (Short: \p{WB=MN}) (15: [,;], U+037E,

U+0589, U+060C..060D, U+066C, U+07F8 ...)

\p{Word_Break: MidNumLet} (Short: \p{WB=MB}) (7: [.],
                   U+2018..2019, U+2024, U+FE52, U+FF07,
                   U+FF0E)

\p{Word_Break: ML}   \p{Word_Break=MidLetter} (9)

\p{Word_Break: MN}   \p{Word_Break=MidNum} (15)

\p{Word_Break: Newline} (Short: \p{WB=NL}) (5: [\cK\f\x85],
                   U+2028..2029)

\p{Word_Break: NL}   \p{Word_Break=Newline} (5)

\p{Word_Break: NU}   \p{Word_Break=Numeric} (651)

\p{Word_Break: Numeric} (Short: \p{WB=NU}) (651: [0-9],
                   U+0660..0669, U+066B, U+06F0..06F9,
                   U+07C0..07C9, U+0966..096F ...)

\p{Word_Break: Other}   (Short: \p{WB=XX}) (1_081_665 plus all
                   above-Unicode code points: [^\n\cK\f\r
                   \x20\"\',.0-9:;A-Z_a-z\x85\xaa\xad\xb5
                   \xb7\xba\xc0-\xd6\xd8-\xf6\xf8-\xff],
                   U+02D8..02DD, U+0375, U+0378..0379,
                   U+0380..0385, U+038B ...)

\p{Word_Break: Regional_Indicator} (Short: \p{WB=RI}) (26:
                   U+1F1E6..1F1FF)

\p{Word_Break: RI}   \p{Word_Break=Regional_Indicator} (26)

\p{Word_Break: Single_Quote} (Short: \p{WB=SQ}) (1: [\'])

\p{Word_Break: SQ}   \p{Word_Break=Single_Quote} (1)

\p{Word_Break: WSegSpace} (Short: \p{WB=WSegSpace}) (14: [\x20],
                   U+1680, U+2000..2006, U+2008..200A,
                   U+205F, U+3000)

\p{Word_Break: XX}   \p{Word_Break=Other} (1_081_665 plus all
                   above-Unicode code points)

\p{Word_Break: ZWJ}   (Short: \p{WB=ZWJ}) (1: U+200D)

\p{WSpace}   \p{White_Space} (= \p{White_Space=Y}) (25)

\p{WSpace: *}   \p{White_Space: *}

\p{XDigit}   \p{XPosixXDigit} (= \p{Hex_Digit=Y}) (44)

\p{XID_Continue}   \p{XID_Continue=Y} (Short: \p{XIDC})
                   (134_415)

\p{XID_Continue: N*}   (Short: \p{XIDC=N}, \P{XIDC}) (979_697
                   plus all above-Unicode code points:
                   [\x00-\x20!\"#\$\%&\'\(\)*+,\-.\/:;<=>?
                   \@\[\\\]\^`\{\|\}~\x7f-\xa9\xab-\xb4
                   \xb6\xb8-\xb9\xbb-\xbf\xd7\xf7],

```
                         U+02C2..02C5, U+02D2..02DF,
                         U+02E5..02EB, U+02ED, U+02EF..02FF ...)
    \p{XID_Continue: Y*}    (Short: \p{XIDC=Y}, \p{XIDC}) (134_415:
                         [0-9A-Z_a-z\xaa\xb5\xb7\xba\xc0-\xd6
                         \xd8-\xf6\xf8-\xff], U+0100..02C1,
                         U+02C6..02D1, U+02E0..02E4, U+02EC,
                         U+02EE ...)
    \p{XID_Start}        \p{XID_Start=Y} (Short: \p{XIDS}) (131_459)
    \p{XID_Start: N*}      (Short: \p{XIDS=N}, \P{XIDS}) (982_653
                         plus all above-Unicode code points:
                         [\x00-\x20!\"#\$\%&\'\(\)*+,\-.\/0-9:;<=
                         >?\@\[\\\]\^_`\{\|\}~\x7f-\xa9\xab-\xb4
                         \xb6-\xb9\xbb-\xbf\xd7\xf7],
                         U+02C2..02C5, U+02D2..02DF,
                         U+02E5..02EB, U+02ED, U+02EF..036F ...)
    \p{XID_Start: Y*}      (Short: \p{XIDS=Y}, \p{XIDS}) (131_459:
                         [A-Za-z\xaa\xb5\xba\xc0-\xd6\xd8-\xf6
                         \xf8-\xff], U+0100..02C1, U+02C6..02D1,
                         U+02E0..02E4, U+02EC, U+02EE ...)
    \p{XIDC}             \p{XID_Continue} (= \p{XID_Continue=Y})
                         (134_415)
    \p{XIDC: *}          \p{XID_Continue: *}
    \p{XIDS}             \p{XID_Start} (= \p{XID_Start=Y}) (131_459)
    \p{XIDS: *}          \p{XID_Start: *}
    \p{Xpeo}             \p{Old_Persian} (= \p{Script_Extensions=
                         Old_Persian}) (NOT \p{Block=
                         Old_Persian}) (50)
    \p{XPerlSpace}       \p{XPosixSpace} (25)
    \p{XPosixAlnum}      Alphabetic and (decimal) Numeric (Short:
                         \p{Alnum}) (133_525: [0-9A-Za-z\xaa\xb5
                         \xba\xc0-\xd6\xd8-\xf6\xf8-\xff],
                         U+0100..02C1, U+02C6..02D1,
                         U+02E0..02E4, U+02EC, U+02EE ...)
    \p{XPosixAlpha}      \p{Alphabetic=Y} (Short: \p{Alpha})
                         (132_875)
    \p{XPosixBlank}      \h, Horizontal white space (Short:
                         \p{Blank}) (18: [\t\x20\xa0], U+1680,
                         U+2000..200A, U+202F, U+205F, U+3000)
    \p{XPosixCntrl}      \p{General_Category=Control} Control
                         characters (Short: \p{Cc}) (65)
```

\p{XPosixDigit}   \p{General_Category=Decimal_Number} [0-9]
                + all other decimal digits (Short:
                \p{Nd}) (650)
\p{XPosixGraph}   Characters that are graphical (Short:
                \p{Graph}) (281_308: [!\"#\$\%&\'
                \(\)*+,\-.\/0-9:;<=>?\@A-Z\[\\\]\^_'a-z
                \{\|\}~\xa1-\xff], U+0100..0377,
                U+037A..037F, U+0384..038A, U+038C,
                U+038E..03A1 ...)
\p{XPosixLower}   \p{Lowercase=Y} (Short: \p{Lower}; /i=
                Cased=Yes) (2344)
\p{XPosixPrint}   Characters that are graphical plus space
                characters (but no controls) (Short:
                \p{Print}) (281_325: [\x20-\x7e\xa0-
                \xff], U+0100..0377, U+037A..037F,
                U+0384..038A, U+038C, U+038E..03A1 ...)
\p{XPosixPunct}   \p{Punct} + ASCII-range \p{Symbol} (807:
                [!\"#\$\%&\'\(\)*+,\-.\/:;<=>?\@\[\\\]
                \^_'\{\|\}~\xa1\xa7\xab\xb6-\xb7\xbb
                \xbf], U+037E, U+0387, U+055A..055F,
                U+0589..058A, U+05BE ...)
\p{XPosixSpace}   \s including beyond ASCII and vertical tab
                (Short: \p{SpacePerl}) (25: [\t\n\cK\f
                \r\x20\x85\xa0], U+1680, U+2000..200A,
                U+2028..2029, U+202F, U+205F ...)
\p{XPosixUpper}   \p{Uppercase=Y} (Short: \p{Upper}; /i=
                Cased=Yes) (1911)
\p{XPosixWord}   \w, including beyond ASCII; = \p{Alnum} +
                \pM + \p{Pc} + \p{Join_Control} (Short:
                \p{Word}) (134_564: [0-9A-Z_a-z\xaa\xb5
                \xba\xc0-\xd6\xd8-\xf6\xf8-\xff],
                U+0100..02C1, U+02C6..02D1,
                U+02E0..02E4, U+02EC, U+02EE ...)
\p{XPosixXDigit}   \p{Hex_Digit=Y} (Short: \p{Hex}) (44)
\p{Xsux}   \p{Cuneiform} (= \p{Script_Extensions=
                Cuneiform}) (NOT \p{Block=Cuneiform})
                (1234)
\p{Yezi}   \p{Yezidi} (= \p{Script_Extensions=
                Yezidi}) (NOT \p{Block=Yezidi}) (60)
\p{Yezidi}   \p{Script_Extensions=Yezidi} (Short:

```
                      \p{Yezi}; NOT \p{Block=Yezidi}) (60)
    \p{Yi}              \p{Script_Extensions=Yi} (1246)
  X \p{Yi_Radicals}     \p{Block=Yi_Radicals} (64)
  X \p{Yi_Syllables}    \p{Block=Yi_Syllables} (1168)
    \p{Yiii}            \p{Yi} (= \p{Script_Extensions=Yi}) (1246)
  X \p{Yijing}          \p{Yijing_Hexagram_Symbols} (= \p{Block=
                      Yijing_Hexagram_Symbols}) (64)
  X \p{Yijing_Hexagram_Symbols} \p{Block=Yijing_Hexagram_Symbols}
                      (Short: \p{InYijing}) (64)
    \p{Z} \pZ          \p{Separator} (= \p{General_Category=
                      Separator}) (19)
    \p{Zanabazar_Square}    \p{Script_Extensions=Zanabazar_Square}
                      (Short: \p{Zanb}; NOT \p{Block=
                      Zanabazar_Square}) (72)
    \p{Zanb}            \p{Zanabazar_Square} (=
                      \p{Script_Extensions=Zanabazar_Square})
                      (NOT \p{Block=Zanabazar_Square}) (72)
    \p{Zinh}            \p{Inherited} (= \p{Script_Extensions=
                      Inherited}) (503)
    \p{Zl}              \p{Line_Separator} (= \p{General_Category=
                      Line_Separator}) (1)
    \p{Zp}              \p{Paragraph_Separator} (=
                      \p{General_Category=
                      Paragraph_Separator}) (1)
    \p{Zs}              \p{Space_Separator} (=
                      \p{General_Category=Space_Separator})
                      (17)
    \p{Zyyy}            \p{Common} (= \p{Script_Extensions=
                      Common}) (7661)
    \p{Zzzz}            \p{Unknown} (= \p{Script_Extensions=
                      Unknown}) (970_188 plus all above-
                      Unicode code points)
```

Legal "\p{}" and "\P{}" constructs that match no characters

Unicode has some property-value pairs that currently don't match anything. This happens generally either because they are obsolete, or they exist for symmetry with other forms, but no language has yet been encoded that uses them. In this version of Unicode, the following match zero code points:

```
\p{Canonical_Combining_Class=Attached_Below_Left}
\p{Canonical_Combining_Class=CCC133}
```

\p{Grapheme_Cluster_Break=E_Base}
\p{Grapheme_Cluster_Break=E_Base_GAZ}
\p{Grapheme_Cluster_Break=E_Modifier}
\p{Grapheme_Cluster_Break=Glue_After_Zwj}
\p{Word_Break=E_Base}
\p{Word_Break=E_Base_GAZ}
\p{Word_Break=E_Modifier}
\p{Word_Break=Glue_After_Zwj}

Properties accessible through Unicode::UCD

The value of any Unicode (not including Perl extensions) character property mentioned above for any single code point is available through "**charprop**()" in Unicode::UCD. "**charprops_all**()" in Unicode::UCD returns the values of all the Unicode properties for a given code point.

Besides these, all the Unicode character properties mentioned above (except for those marked as for internal use by Perl) are also accessible by "**prop_invlist**()" in Unicode::UCD.

Due to their nature, not all Unicode character properties are suitable for regular expression matches, nor "prop_invlist()". The remaining non-provisional, non-internal ones are accessible via "**prop_invmap**()" in Unicode::UCD (except for those that this Perl installation hasn't included; see below for which those are).

For compatibility with other parts of Perl, all the single forms given in the table in the section above are recognized. BUT, there are some ambiguities between some Perl extensions and the Unicode properties, all of which are silently resolved in favor of the official Unicode property. To avoid surprises, you should only use "prop_invmap()" for forms listed in the table below, which omits the non-recommended ones. The affected forms are the Perl single form equivalents of Unicode properties, such as "\p{sc}" being a single-form equivalent of "\p{gc=sc}", which is treated by "prop_invmap()" as the "Script" property, whose short name is "sc". The table indicates the current ambiguities in the INFO column, beginning with the word "NOT".

The standard Unicode properties listed below are documented in <http://www.unicode.org/reports/tr44/>; Perl_Decimal_Digit is documented in "**prop_invmap**()" in Unicode::UCD. The other Perl extensions are in "Other Properties" in perlunicode;

The first column in the table is a name for the property; the second column is an alternative name, if any, plus possibly some annotations. The alternative name is the property's full name, unless that would simply repeat the first column, in which case the second column indicates the property's short name (if different). The annotations are given only in the entry for the full name. The annotations for binary properties include a list of the first few ranges that the property matches. To avoid any

ambiguity, the SPACE character is represented as "\x20".

If a property is obsolete, etc, the entry will be flagged with the same characters used in the table in the section above, like **D** or **S**.

```
NAME                INFO

Age
AHex              ASCII_Hex_Digit
All                (Perl extension).  All code points,
                   including those above Unicode.  Same as
                   qr/./s.  U+0000..infinity
Alnum              XPosixAlnum.  (Perl extension)
Alpha             Alphabetic
Alphabetic         (Short: Alpha).  [A-Za-z\xaa\xb5\xba\xc0-
                   \xd6\xd8-\xf6\xf8-\xff], U+0100..02C1,
                   U+02C6..02D1, U+02E0..02E4, U+02EC, U+02EE
                   ...
Any               (Perl extension).  All Unicode code
                   points.  U+0000..10FFFF
ASCII              Block=Basic_Latin.  (Perl extension).
                   [\x00-\x7f]
ASCII_Hex_Digit    (Short: AHex).  [0-9A-Fa-f]
Assigned           (Perl extension).  All assigned code
                   points.  U+0000..0377, U+037A..037F,
                   U+0384..038A, U+038C, U+038E..03A1,
                   U+03A3..052F ...
Bc                Bidi_Class
Bidi_C            Bidi_Control
Bidi_Class         (Short: bc)
Bidi_Control       (Short: Bidi_C).  U+061C, U+200E..200F,
                   U+202A..202E, U+2066..2069
Bidi_M            Bidi_Mirrored
Bidi_Mirrored      (Short: Bidi_M).  [\(\)<>\[\]\{\}\xab
                   \xbb], U+0F3A..0F3D, U+169B..169C,
                   U+2039..203A, U+2045..2046, U+207D..207E
                   ...
Bidi_Mirroring_Glyph   (Short: bmg)
Bidi_Paired_Bracket    (Short: bpb)
Bidi_Paired_Bracket_Type (Short: bpt)
```

Blank   XPosixBlank. (Perl extension)

Blk   Block

Block   (Short: blk)

Bmg   Bidi_Mirroring_Glyph

Bpb   Bidi_Paired_Bracket

Bpt   Bidi_Paired_Bracket_Type

Canonical_Combining_Class (Short: ccc)

Case_Folding   (Short: cf)

Case_Ignorable   (Short: CI). [\'.:\^'\xa8\xad\xaf\xb4
                 \xb7-\xb8], U+02B0..036F, U+0374..0375,
                 U+037A, U+0384..0385, U+0387 ...

Cased   [A-Za-z\xaa\xb5\xba\xc0-\xd6\xd8-\xf6\xf8-
                 \xff], U+0100..01BA, U+01BC..01BF,
                 U+01C4..0293, U+0295..02B8, U+02C0..02C1

                 ...

Category   General_Category

Ccc   Canonical_Combining_Class

CE   Composition_Exclusion

Cf   Case_Folding; NOT 'cf' meaning
                 'General_Category=Format'

Changes_When_Casefolded (Short: CWCF). [A-Z\xb5\xc0-\xd6\xd8-
                 \xdf], U+0100, U+0102, U+0104, U+0106,
                 U+0108 ...

Changes_When_Casemapped (Short: CWCM). [A-Za-z\xb5\xc0-\xd6\xd8-
                 \xf6\xf8-\xff], U+0100..0137,
                 U+0139..018C, U+018E..019A, U+019C..01A9,
                 U+01AC..01B9 ...

Changes_When_Lowercased (Short: CWL). [A-Z\xc0-\xd6\xd8-\xde],
                 U+0100, U+0102, U+0104, U+0106, U+0108 ...

Changes_When_NFKC_Casefolded (Short: CWKCF). [A-Z\xa0\xa8\xaa
                 \xad\xaf\xb2-\xb5\xb8-\xba\xbc-\xbe\xc0-
                 \xd6\xd8-\xdf], U+0100, U+0102, U+0104,
                 U+0106, U+0108 ...

Changes_When_Titlecased (Short: CWT). [a-z\xb5\xdf-\xf6\xf8-
                 \xff], U+0101, U+0103, U+0105, U+0107,
                 U+0109 ...

Changes_When_Uppercased (Short: CWU). [a-z\xb5\xdf-\xf6\xf8-
                 \xff], U+0101, U+0103, U+0105, U+0107,
                 U+0109 ...

CI   Case_Ignorable

```
Cntrl              XPosixCntrl (=General_Category=Control).
                   (Perl extension)
Comp_Ex            Full_Composition_Exclusion
Composition_Exclusion   (Short: CE).  U+0958..095F, U+09DC..09DD,
                   U+09DF, U+0A33, U+0A36, U+0A59..0A5B ...
CWCF               Changes_When_Casefolded
CWCM               Changes_When_Casemapped
CWKCF              Changes_When_NFKC_Casefolded
CWL                Changes_When_Lowercased
CWT                Changes_When_Titlecased
CWU                Changes_When_Uppercased
Dash               [\-], U+058A, U+05BE, U+1400, U+1806,
                   U+2010..2015 ...
Decomposition_Mapping   (Short: dm)
Decomposition_Type     (Short: dt)
Default_Ignorable_Code_Point (Short: DI).  [\xad], U+034F, U+061C,
                   U+115F..1160, U+17B4..17B5, U+180B..180E

                   ...
Dep                Deprecated
Deprecated             (Short: Dep).  U+0149, U+0673, U+0F77,
                   U+0F79, U+17A3..17A4, U+206A..206F ...
DI                 Default_Ignorable_Code_Point
Dia                Diacritic
Diacritic              (Short: Dia).  [\^`\xa8\xaf\xb4\xb7-\xb8],
                   U+02B0..034E, U+0350..0357, U+035D..0362,
                   U+0374..0375, U+037A ...
Digit              XPosixDigit (=General_Category=
                   Decimal_Number).  (Perl extension)
Dm                 Decomposition_Mapping
Dt                 Decomposition_Type
Ea                 East_Asian_Width
East_Asian_Width       (Short: ea)
EBase              Emoji_Modifier_Base
EComp              Emoji_Component
EMod               Emoji_Modifier
Emoji              [#*0-9\xa9\xae], U+203C, U+2049, U+2122,
                   U+2139, U+2194..2199 ...
Emoji_Component        (Short: EComp).  [#*0-9], U+200D, U+20E3,
                   U+FE0F, U+1F1E6..1F1FF, U+1F3FB..1F3FF ...
Emoji_Modifier         (Short: EMod).  U+1F3FB..1F3FF
```

```
Emoji_Modifier_Base    (Short: EBase).  U+261D, U+26F9,
               U+270A..270D, U+1F385, U+1F3C2..1F3C4,
               U+1F3C7 ...
Emoji_Presentation     (Short: EPres).  U+231A..231B,
               U+23E9..23EC, U+23F0, U+23F3,
               U+25FD..25FE, U+2614..2615 ...
EPres             Emoji_Presentation
EqUIdeo           Equivalent_Unified_Ideograph
Equivalent_Unified_Ideograph (Short: EqUIdeo)
Ext             Extender
Extended_Pictographic   (Short: ExtPict).  [\xa9\xae], U+203C,
               U+2049, U+2122, U+2139, U+2194..2199 ...
Extender          (Short: Ext).  [\xb7], U+02D0..02D1,
               U+0640, U+07FA, U+0B55, U+0E46 ...
ExtPict           Extended_Pictographic
Full_Composition_Exclusion (Short: Comp_Ex).  U+0340..0341,
               U+0343..0344, U+0374, U+037E, U+0387,
               U+0958..095F ...
Gc             General_Category
GCB             Grapheme_Cluster_Break
General_Category      (Short: gc)
Gr_Base          Grapheme_Base
Gr_Ext           Grapheme_Extend
Graph           XPosixGraph.  (Perl extension)
Grapheme_Base       (Short: Gr_Base).  [\x20-\x7e\xa0-\xac
               \xae-\xff], U+0100..02FF, U+0370..0377,
               U+037A..037F, U+0384..038A, U+038C ...
Grapheme_Cluster_Break  (Short: GCB)
Grapheme_Extend      (Short: Gr_Ext).  U+0300..036F,
               U+0483..0489, U+0591..05BD, U+05BF,
               U+05C1..05C2, U+05C4..05C5 ...
Hangul_Syllable_Type   (Short: hst)
Hex             Hex_Digit
Hex_Digit          (Short: Hex).  [0-9A-Fa-f], U+FF10..FF19,
               U+FF21..FF26, U+FF41..FF46
HorizSpace         XPosixBlank.  (Perl extension)
Hst             Hangul_Syllable_Type
D Hyphen           [\-\xad], U+058A, U+1806, U+2010..2011,
               U+2E17, U+30FB ...  Supplanted by
               Line_Break property values; see
```

```
                www.unicode.org/reports/tr14
    ID_Continue          (Short: IDC).  [0-9A-Z_a-z\xaa\xb5\xb7
                \xba\xc0-\xd6\xd8-\xf6\xf8-\xff],
                U+0100..02C1, U+02C6..02D1, U+02E0..02E4,
                U+02EC, U+02EE ...
    ID_Start            (Short: IDS).  [A-Za-z\xaa\xb5\xba\xc0-
                \xd6\xd8-\xf6\xf8-\xff], U+0100..02C1,
                U+02C6..02D1, U+02E0..02E4, U+02EC, U+02EE

                ...
    IDC            ID_Continue
    Identifier_Status
    Identifier_Type
    Ideo            Ideographic
    Ideographic         (Short: Ideo).  U+3006..3007,
                U+3021..3029, U+3038..303A, U+3400..4DBF,
                U+4E00..9FFC, U+F900..FA6D ...
    IDS            ID_Start
    IDS_Binary_Operator     (Short: IDSB).  U+2FF0..2FF1, U+2FF4..2FFB
    IDS_Trinary_Operator    (Short: IDST).  U+2FF2..2FF3
    IDSB            IDS_Binary_Operator
    IDST            IDS_Trinary_Operator
    In            Present_In.  (Perl extension)
    Indic_Positional_Category (Short: InPC)
    Indic_Syllabic_Category (Short: InSC)
    InPC            Indic_Positional_Category
    InSC            Indic_Syllabic_Category
    Isc            ISO_Comment; NOT 'isc' meaning
                'General_Category=Other'
    ISO_Comment         (Short: isc)
    Jg            Joining_Group
    Join_C            Join_Control
    Join_Control         (Short: Join_C).  U+200C..200D
    Joining_Group          (Short: jg)
    Joining_Type          (Short: jt)
    Jt            Joining_Type
    Lb             Line_Break
    Lc             Lowercase_Mapping; NOT 'lc' meaning
                'General_Category=Cased_Letter'
    Line_Break          (Short: lb)
    LOE            Logical_Order_Exception
```

Logical_Order_Exception (Short: LOE). U+0E40..0E44, U+0EC0..0EC4,
            U+19B5..19B7, U+19BA, U+AAB5..AAB6, U+AAB9

            ...
Lower   Lowercase
Lowercase   (Short: Lower). [a-z\xaa\xb5\xba\xdf-
            \xf6\xf8-\xff], U+0101, U+0103, U+0105,
            U+0107, U+0109 ...
Lowercase_Mapping   (Short: lc)
Math   [+<=>\^\|~\xac\xb1\xd7\xf7], U+03D0..03D2,
            U+03D5, U+03F0..03F1, U+03F4..03F6,
            U+0606..0608 ...
Na   Name
Na1   Unicode_1_Name
Name   (Short: na)
Name_Alias
NChar   Noncharacter_Code_Point
NFC_QC   NFC_Quick_Check
NFC_Quick_Check   (Short: NFC_QC)
NFD_QC   NFD_Quick_Check
NFD_Quick_Check   (Short: NFD_QC)
NFKC_Casefold   (Short: NFKC_CF)
NFKC_CF   NFKC_Casefold
NFKC_QC   NFKC_Quick_Check
NFKC_Quick_Check   (Short: NFKC_QC)
NFKD_QC   NFKD_Quick_Check
NFKD_Quick_Check   (Short: NFKD_QC)
Noncharacter_Code_Point (Short: NChar). U+FDD0..FDEF,
            U+FFFE..FFFF, U+1FFFE..1FFFF,
            U+2FFFE..2FFFF, U+3FFFE..3FFFF,
            U+4FFFE..4FFFF ...
Nt   Numeric_Type
Numeric_Type   (Short: nt)
Numeric_Value   (Short: nv)
Nv   Numeric_Value
Pat_Syn   Pattern_Syntax
Pat_WS   Pattern_White_Space
Pattern_Syntax   (Short: Pat_Syn). [!\"#\$\%&\'\(\)*+,\-.
            \/:;<=>?\@\[\\\]\^'\{\|\}~\xa1-\xa7\xa9
            \xab-\xac\xae\xb0-\xb1\xb6\xbb\xbf\xd7
            \xf7], U+2010..2027, U+2030..203E,

U+2041..2053, U+2055..205E, U+2190..245F

...

Pattern_White_Space   (Short: Pat_WS). [\t\n\cK\f\r\x20\x85],
                       U+200E..200F, U+2028..2029
PCM   Prepended_Concatenation_Mark
Perl_Decimal_Digit   (Perl extension)
PerlSpace   PosixSpace. (Perl extension)
PerlWord   PosixWord. (Perl extension)
PosixAlnum   (Perl extension). [0-9A-Za-z]
PosixAlpha   (Perl extension). [A-Za-z]
PosixBlank   (Perl extension). [\t\x20]
PosixCntrl   (Perl extension). ASCII control
                    characters. ACK, BEL, BS, CAN, CR, DC1,
                    DC2, DC3, DC4, DEL, DLE, ENQ, EOM, EOT,
                    ESC, ETB, ETX, FF, FS, GS, HT, LF, NAK,
                    NUL, RS, SI, SO, SOH, STX, SUB, SYN, US, VT
PosixDigit   (Perl extension). [0-9]
PosixGraph   (Perl extension). [!\"#\$\%&\'\(\)*+,\-.
                    \/0-9:;<=>?\@A-Z\[\\\]\^_'a-z\{\|\}~]
PosixLower   (Perl extension). [a-z]
PosixPrint   (Perl extension). [\x20-\x7e]
PosixPunct   (Perl extension). [!\"#\$\%&\'\(\)*+,\-.
                    \/:;<=>?\@\[\\\]\^_'\{\|\}~]
PosixSpace   (Perl extension). [\t\n\cK\f\r\x20]
PosixUpper   (Perl extension). [A-Z]
PosixWord   (Perl extension). \w, restricted to
                    ASCII. [0-9A-Z_a-z]
PosixXDigit   ASCII_Hex_Digit. (Perl extension).
                    [0-9A-Fa-f]
Prepended_Concatenation_Mark (Short: PCM). U+0600..0605, U+06DD,
                    U+070F, U+08E2, U+110BD, U+110CD
Present_In   (Short: In). (Perl extension)
Print   XPosixPrint. (Perl extension)
Punct   General_Category=Punctuation. (Perl
                    extension). [!\"#\%&\'\(\)*,\-.\/:;?\@
                    \[\\\]_\{\}\xa1\xa7\xab\xb6-\xb7\xbb\xbf],
                    U+037E, U+0387, U+055A..055F,
                    U+0589..058A, U+05BE ...
QMark   Quotation_Mark
Quotation_Mark   (Short: QMark). [\"\'\xab\xbb],

```
                    U+2018..201F, U+2039..203A, U+2E42,
                    U+300C..300F, U+301D..301F ...
       Radical           U+2E80..2E99, U+2E9B..2EF3, U+2F00..2FD5
       Regional_Indicator    (Short: RI).  U+1F1E6..1F1FF
       RI                Regional_Indicator
       SB                Sentence_Break
       Sc                Script; NOT 'sc' meaning
                    'General_Category=Currency_Symbol'
       Scf               Simple_Case_Folding
       Script            (Short: sc)
       Script_Extensions     (Short: scx)
       Scx               Script_Extensions
       SD                Soft_Dotted
       Sentence_Break        (Short: SB)
       Sentence_Terminal     (Short: STerm).  [!.?], U+0589,
                    U+061E..061F, U+06D4, U+0700..0702, U+07F9
                    ...
       Sfc               Simple_Case_Folding
       Simple_Case_Folding    (Short: scf)
       Simple_Lowercase_Mapping (Short: slc)
       Simple_Titlecase_Mapping (Short: stc)
       Simple_Uppercase_Mapping (Short: suc)
       Slc               Simple_Lowercase_Mapping
       Soft_Dotted           (Short: SD).  [i-j], U+012F, U+0249,
                    U+0268, U+029D, U+02B2 ...
       Space             White_Space
       SpacePerl             XPosixSpace.  (Perl extension)
       Stc               Simple_Titlecase_Mapping
       STerm             Sentence_Terminal
       Suc               Simple_Uppercase_Mapping
       Tc                Titlecase_Mapping
       Term              Terminal_Punctuation
       Terminal_Punctuation    (Short: Term).  [!,.:;?], U+037E, U+0387,
                    U+0589, U+05C3, U+060C ...
       Title             Titlecase.  (Perl extension)
       Titlecase             (Short: Title).  (Perl extension).  (=
                    \p{Gc=Lt}).  U+01C5, U+01C8, U+01CB,
                    U+01F2, U+1F88..1F8F, U+1F98..1F9F ...
       Titlecase_Mapping     (Short: tc)
       Uc                Uppercase_Mapping
```

    UIdeo   Unified_Ideograph
    Unicode   Any. (Perl extension)
    Unicode_1_Name   (Short: na1)
    Unified_Ideograph   (Short: UIdeo). U+3400..4DBF,
                U+4E00..9FFC, U+FA0E..FA0F, U+FA11,
                U+FA13..FA14, U+FA1F ...
    Upper   Uppercase
    Uppercase   (Short: Upper). [A-Z\xc0-\xd6\xd8-\xde],
                U+0100, U+0102, U+0104, U+0106, U+0108 ...
    Uppercase_Mapping   (Short: uc)
    Variation_Selector   (Short: VS). U+180B..180D, U+FE00..FE0F,
                U+E0100..E01EF
    Vertical_Orientation   (Short: vo)
    VertSpace   (Perl extension). \v. [\n\cK\f\r\x85],
                U+2028..2029
    Vo   Vertical_Orientation
    VS   Variation_Selector
    WB   Word_Break
    White_Space   (Short: WSpace). [\t\n\cK\f\r\x20\x85
                \xa0], U+1680, U+2000..200A, U+2028..2029,
                U+202F, U+205F ...
    Word   XPosixWord. (Perl extension)
    Word_Break   (Short: WB)
    WSpace   White_Space
    XDigit   XPosixXDigit (=Hex_Digit). (Perl
                extension)
    XID_Continue   (Short: XIDC). [0-9A-Z_a-z\xaa\xb5\xb7
                \xba\xc0-\xd6\xd8-\xf6\xf8-\xff],
                U+0100..02C1, U+02C6..02D1, U+02E0..02E4,
                U+02EC, U+02EE ...
    XID_Start   (Short: XIDS). [A-Za-z\xaa\xb5\xba\xc0-
                \xd6\xd8-\xf6\xf8-\xff], U+0100..02C1,
                U+02C6..02D1, U+02E0..02E4, U+02EC, U+02EE
                ...
    XIDC   XID_Continue
    XIDS   XID_Start
    XPerlSpace   XPosixSpace. (Perl extension)
    XPosixAlnum   (Short: Alnum). (Perl extension).
                Alphabetic and (decimal) Numeric. [0-9A-
                Za-z\xaa\xb5\xba\xc0-\xd6\xd8-\xf6\xf8-

                \xff], U+0100..02C1, U+02C6..02D1,
                U+02E0..02E4, U+02EC, U+02EE ...

XPosixAlpha   Alphabetic. (Perl extension). [A-Za-z
                \xaa\xb5\xba\xc0-\xd6\xd8-\xf6\xf8-\xff],
                U+0100..02C1, U+02C6..02D1, U+02E0..02E4,
                U+02EC, U+02EE ...

XPosixBlank   (Short: Blank). (Perl extension). \h,
                Horizontal white space. [\t\x20\xa0],
                U+1680, U+2000..200A, U+202F, U+205F,
                U+3000

XPosixCntrl   General_Category=Control (Short: Cntrl).
                (Perl extension). Control characters.
                [\x00-\x1f\x7f-\x9f]

XPosixDigit   General_Category=Decimal_Number (Short:
                Digit). (Perl extension). [0-9] + all
                other decimal digits. [0-9],
                U+0660..0669, U+06F0..06F9, U+07C0..07C9,
                U+0966..096F, U+09E6..09EF ...

XPosixGraph   (Short: Graph). (Perl extension).
                Characters that are graphical. [!\"#\$
                \%&\'\(\)*+,\-.\/0-9:;<=>?\@A-Z\[\\\]
                \^_'a-z\{\|\}~\xa1-\xff], U+0100..0377,
                U+037A..037F, U+0384..038A, U+038C,
                U+038E..03A1 ...

XPosixLower   Lowercase. (Perl extension). [a-z\xaa
                \xb5\xba\xdf-\xf6\xf8-\xff], U+0101,
                U+0103, U+0105, U+0107, U+0109 ...

XPosixPrint   (Short: Print). (Perl extension).
                Characters that are graphical plus space
                characters (but no controls). [\x20-\x7e
                \xa0-\xff], U+0100..0377, U+037A..037F,
                U+0384..038A, U+038C, U+038E..03A1 ...

XPosixPunct   (Perl extension). \p{Punct} + ASCII-range
                \p{Symbol}. [!\"#\$\%&\'\(\)*+,\-.\/:;<=
                >?\@\[\\\]\^_'\{\|\}~\xa1\xa7\xab\xb6-
                \xb7\xbb\xbf], U+037E, U+0387,
                U+055A..055F, U+0589..058A, U+05BE ...

XPosixSpace   (Perl extension). \s including beyond
                ASCII and vertical tab. [\t\n\cK\f\r\x20
                \x85\xa0], U+1680, U+2000..200A,

	U+2028..2029, U+202F, U+205F ...
XPosixUpper	Uppercase. (Perl extension). [A-Z\xc0-\xd6\xd8-\xde], U+0100, U+0102, U+0104, U+0106, U+0108 ...
XPosixWord	(Short: Word). (Perl extension). \w, including beyond ASCII; = \p{Alnum} + \pM + \p{Pc} + \p{Join_Control}. [0-9A-Z_a-z\xaa\xb5\xba\xc0-\xd6\xd8-\xf6\xf8-\xff], U+0100..02C1, U+02C6..02D1, U+02E0..02E4, U+02EC, U+02EE ...
XPosixXDigit	Hex_Digit (Short: XDigit). (Perl extension). [0-9A-Fa-f], U+FF10..FF19, U+FF21..FF26, U+FF41..FF46

Properties accessible through other means

Certain properties are accessible also via core function calls. These are:

Lowercase_Mapping	lc() and lcfirst()
Titlecase_Mapping	ucfirst()
Uppercase_Mapping	uc()

Also, Case_Folding is accessible through the "/i" modifier in regular expressions, the "\F" transliteration escape, and the "fc" operator.

Besides being able to say "\p{Name=...}", the Name and Name_Aliases properties are accessible through the "\N{}" interpolation in double-quoted strings and regular expressions; and functions "charnames::viacode()", "charnames::vianame()", and "charnames::string_vianame()" (which require a "use charnames ();" to be specified.

Finally, most properties related to decomposition are accessible via Unicode::Normalize.

Unicode character properties that are NOT accepted by Perl

Perl will generate an error for a few character properties in Unicode when used in a regular expression. The non-Unihan ones are listed below, with the reasons they are not accepted, perhaps with work-arounds. The short names for the properties are listed enclosed in (parentheses). As described after the list, an installation can change the defaults and choose to accept any of these. The list is machine generated based on the choices made for the installation that generated this document.

Expands_On_NFC (XO_NFC)
Expands_On_NFD (XO_NFD)

Expands_On_NFKC (XO_NFKC)
Expands_On_NFKD (XO_NFKD)
   Deprecated by Unicode. These are characters that expand to more than one character in the
   specified normalization form, but whether they actually take up more bytes or not depends on the
   encoding being used. For example, a UTF-8 encoded character may expand to a different number
   of bytes than a UTF-32 encoded character.

Grapheme_Link (Gr_Link)
   Duplicates ccc=vr (Canonical_Combining_Class=Virama)

Jamo_Short_Name (JSN)
Other_Alphabetic (OAlpha)
Other_Default_Ignorable_Code_Point (ODI)
Other_Grapheme_Extend (OGr_Ext)
Other_ID_Continue (OIDC)
Other_ID_Start (OIDS)
Other_Lowercase (OLower)
Other_Math (OMath)
Other_Uppercase (OUpper)
   Used by Unicode internally for generating other properties and not intended to be used stand-
   alone

Script=Katakana_Or_Hiragana (sc=Hrkt)
   Obsolete. All code points previously matched by this have been moved to "Script=Common".
   Consider instead using "Script_Extensions=Katakana" or "Script_Extensions=Hiragana" (or both)

Script_Extensions=Katakana_Or_Hiragana (scx=Hrkt)
   All code points that would be matched by this are matched by either
   "Script_Extensions=Katakana" or "Script_Extensions=Hiragana"

An installation can choose to allow any of these to be matched by downloading the Unicode database
from <http://www.unicode.org/Public/> to $Config{privlib}/*unicore/* in the Perl source tree, changing
the controlling lists contained in the program $Config{privlib}/*unicore/mktables* and then re-compiling
and installing. (%Config is available from the Config module).

Also, perl can be recompiled to operate on an earlier version of the Unicode standard. Further
information is at $Config{privlib}/*unicore/README.perl*.

Other information in the Unicode data base
   The Unicode data base is delivered in two different formats. The XML version is valid for more

modern Unicode releases. The other version is a collection of files. The two are intended to give equivalent information. Perl uses the older form; this allows you to recompile Perl to use early Unicode releases.

The only non-character property that Perl currently supports is Named Sequences, in which a sequence of code points is given a name and generally treated as a single entity. (Perl supports these via the "\N{...}" double-quotish construct, "charnames::string_vianame(name)" in charnames, and "**namedseq()**" in Unicode::UCD.

Below is a list of the files in the Unicode data base that Perl doesn't currently use, along with very brief descriptions of their purposes. Some of the names of the files have been shortened from those that Unicode uses, in order to allow them to be distinguishable from similarly named files on file systems for which only the first 8 characters of a name are significant.

auxiliary/GraphemeBreakTest.html
auxiliary/LineBreakTest.html
auxiliary/SentenceBreakTest.html
auxiliary/WordBreakTest.html
      Documentation of validation Tests

BidiCharacterTest.txt
BidiTest.txt
NormTest.txt
      Validation Tests

CJKRadicals.txt
      Maps the kRSUnicode property values to corresponding code points

emoji/ReadMe.txt
ReadMe.txt
      Documentation

EmojiSources.txt
      Maps certain Unicode code points to their legacy Japanese cell-phone values

extracted/DName.txt
      This file adds no new information not already present in other files

Index.txt
      Alphabetical index of Unicode characters

NamedSqProv.txt
    Named sequences proposed for inclusion in a later version of the Unicode Standard; if you need
    them now, you can append this file to *NamedSequences.txt* and recompile perl

NamesList.html
    Describes the format and contents of *NamesList.txt*

NamesList.txt
    Annotated list of characters

NormalizationCorrections.txt
    Documentation of corrections already incorporated into the Unicode data base

NushuSources.txt
    Specifies source material for Nushu characters

StandardizedVariants.html
    Obsoleted as of Unicode 9.0, but previously provided a visual display of the standard variant
    sequences derived from *StandardizedVariants.txt*.

StandardizedVariants.txt
    Certain glyph variations for character display are standardized. This lists the non-Unihan ones;
    the Unihan ones are also not used by Perl, and are in a separate Unicode data base
    <http://www.unicode.org/ivd>

TangutSources.txt
    Specifies source mappings for Tangut ideographs and components. This data file also includes
    informative radical-stroke values that are used internally by Unicode

USourceData.txt
    Documentation of status and cross reference of proposals for encoding by Unicode of Unihan
    characters

USourceGlyphs.pdf
    Pictures of the characters in *USourceData.txt*

SEE ALSO
<http://www.unicode.org/reports/tr44/>

perlrecharclass

perlunicode