

# **GNU Libidn API Reference Manual**

**COLLABORATORS**

	<i>TITLE :</i> GNU Libidn API Reference Manual		
<i>ACTION</i>	<i>NAME</i>	<i>DATE</i>	<i>SIGNATURE</i>
WRITTEN BY		May 15, 2021	

**REVISION HISTORY**

NUMBER	DATE	DESCRIPTION	NAME

# Contents

<b>1</b>	<b>GNU Libidn API Reference Manual</b>	<b>1</b>
1.1	idna . . . . .	2
1.2	stringprep . . . . .	15
1.3	punycode . . . . .	25
1.4	pr29 . . . . .	31
1.5	tld . . . . .	35
1.6	idn-free . . . . .	35
<b>2</b>	<b>Index</b>	<b>37</b>

# List of Figures

1.1	Components of Libidn . . . . .	2
-----	--------------------------------	---

## Chapter 1

# GNU Libidn API Reference Manual

GNU Libidn is a fully documented implementation of the Stringprep, Punycode and IDNA specifications. Libidn's purpose is to encode and decode internationalized domain name strings. There are native C, C# and Java libraries.

The C library contains a generic Stringprep implementation. Profiles for Nameprep, iSCSI, SASL, XMPP and Kerberos V5 are included. Punycode and ASCII Compatible Encoding (ACE) via IDNA are supported. A mechanism to define Top-Level Domain (TLD) specific validation tables, and to compare strings against those tables, is included. Default tables for some TLDs are also included.

The Stringprep API consists of two main functions, one for converting data from the system's native representation into UTF-8, and one function to perform the Stringprep processing. Adding a new Stringprep profile for your application within the API is straightforward. The Punycode API consists of one encoding function and one decoding function. The IDNA API consists of the ToASCII and ToUnicode functions, as well as an high-level interface for converting entire domain names to and from the ACE encoded form. The TLD API consists of one set of functions to extract the TLD name from a domain string, one set of functions to locate the proper TLD table to use based on the TLD name, and core functions to validate a string against a TLD table, and some utility wrappers to perform all the steps in one call.

The library is used by, e.g., GNU SASL and Shishi to process user names and passwords. Libidn can be built into GNU Libc to enable a new system-wide getaddrinfo flag for IDN processing.

Libidn is developed for the GNU/Linux system, but runs on over 20 Unix platforms (including Solaris, IRIX, AIX, and Tru64) and Windows. The library is written in C and (parts of) the API is also accessible from C++, Emacs Lisp, Python and Java. A native Java and C# port is included.

Also included is a command line tool, several self tests, code examples, and more.

The internal layout of the library, and how your application interact with the various parts of the library, are shown in Figure 1.1.

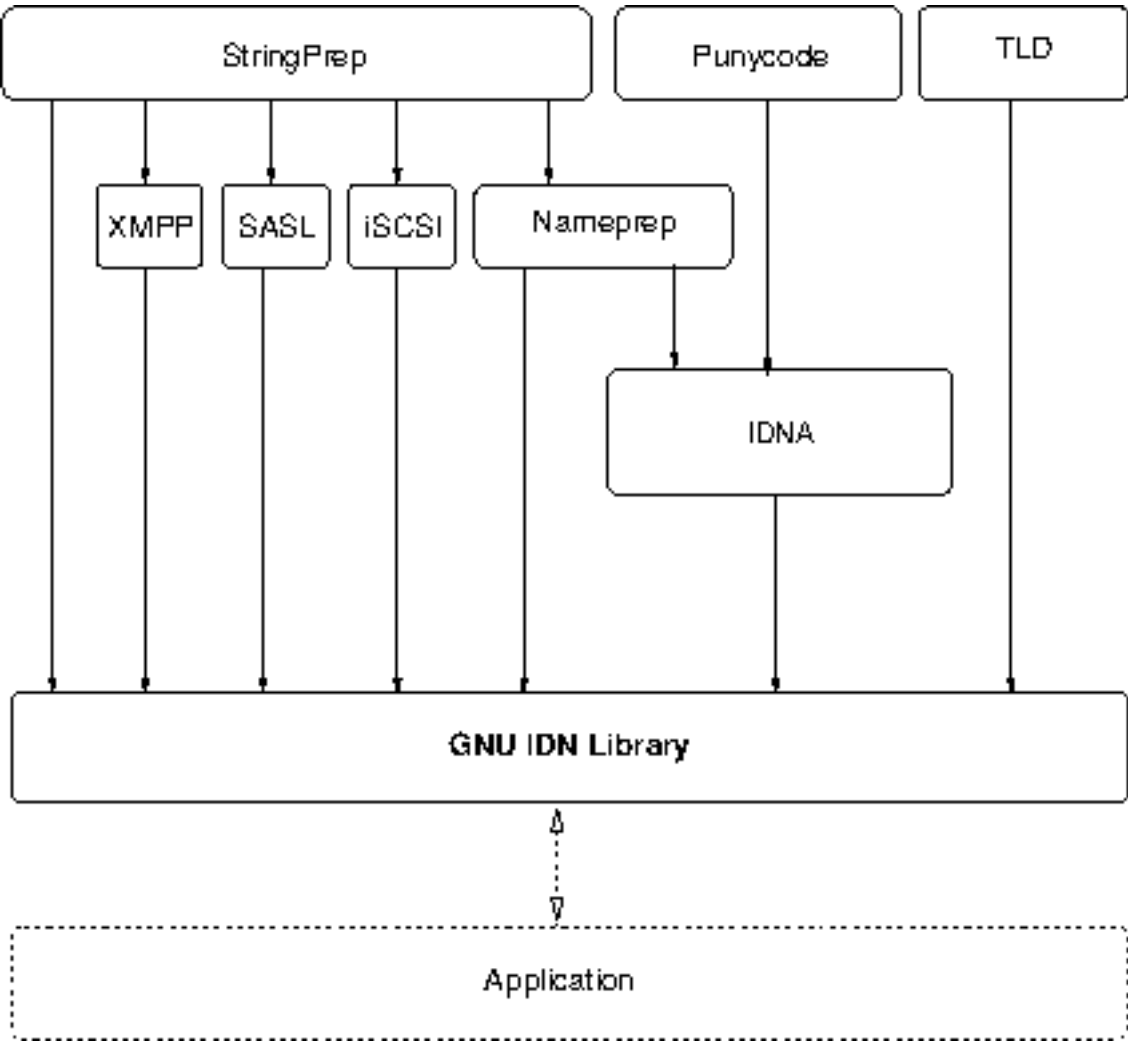


Figure 1.1: Components of Libidn

1.1 idna

idna —

Functions

<code>const char *</code>	<code>idna_strerror ()</code>
<code>int</code>	<code>idna_to_ascii_4i ()</code>
<code>int</code>	<code>idna_to_unicode_44i ()</code>
<code>int</code>	<code>idna_to_ascii_4z ()</code>
<code>int</code>	<code>idna_to_ascii_8z ()</code>
<code>int</code>	<code>idna_to_ascii_lz ()</code>
<code>int</code>	<code>idna_to_unicode_4z4z ()</code>
<code>int</code>	<code>idna_to_unicode_8z4z ()</code>
<code>int</code>	<code>idna_to_unicode_8z8z ()</code>
<code>int</code>	<code>idna_to_unicode_8z1z ()</code>
<code>int</code>	<code>idna_to_unicode_l1zlz ()</code>

Types and Values

#define	IDNAPI
enum	Idna_rc
enum	Idna_flags
#define	IDNA_ACE_PREFIX

Description

Functions

idna\_strerror ()

```
const char~*
idna_strerror (Idna_rc rc);
```

Convert a return code integer to a text string. This string can be used to output a diagnostic message to the user.

IDNA\_SUCCESS: Successful operation. This value is guaranteed to always be zero, the remaining ones are only guaranteed to hold non-zero values, for logical comparison purposes. IDNA\_STRINGPREP\_ERROR: Error during string preparation. IDNA\_PUNYCODE\_ERROR: Error during punycode operation. IDNA\_CONTAINS\_NON\_LDH: For IDNA\_USE\_STD3\_ASCII\_RULES, indicate that the string contains non-LDH ASCII characters. IDNA\_CONTAINS\_MINUS: For IDNA\_USE\_STD3\_ASCII\_RULES, indicate that the string contains a leading or trailing hyphen-minus (U+002D). IDNA\_INVALID\_LENGTH: The final output string is not within the (inclusive) range 1 to 63 characters. IDNA\_NO\_ACE\_PREFIX: The string does not contain the ACE prefix (for ToUnicode). IDNA\_ROUNDTRIP\_VERIFY\_ERROR: The ToASCII operation on output string does not equal the input. IDNA\_CONTAINS\_ACE\_PREFIX: The input contains the ACE prefix (for ToASCII). IDNA\_ICONV\_ERROR: Character encoding conversion error. IDNA\_MALLOC\_ERROR: Could not allocate buffer (this is typically a fatal error). IDNA\_DLOPEN\_ERROR: Could not dlopen the libcidn DSO (only used internally in libc).

Parameters

rc | an Idna\_rc return code. |

Returns

Returns a pointer to a statically allocated string containing a description of the error with the return code *rc*.

idna\_to\_ascii\_4i ()

```
int
idna_to_ascii_4i (const uint32_t *in,
                 size_t inlen,
                 char *out,
                 int flags);
```

The ToASCII operation takes a sequence of Unicode code points that make up one domain label and transforms it into a sequence of code points in the ASCII range (0..7F). If ToASCII succeeds, the original sequence and the resulting sequence are equivalent labels.

It is important to note that the ToASCII operation can fail. ToASCII fails if any step of it fails. If any step of the ToASCII operation fails on any label in a domain name, that domain name MUST NOT be used as an internationalized domain name. The method for deadling with this failure is application-specific.

The inputs to ToASCII are a sequence of code points, the AllowUnassigned flag, and the UseSTD3ASCIIRules flag. The output of ToASCII is either a sequence of ASCII code points or a failure condition.

ToASCII never alters a sequence of code points that are all in the ASCII range to begin with (although it could fail). Applying the ToASCII operation multiple times has exactly the same effect as applying it just once.

Parameters

in	input array with unicode code points.	
inlen	length of input array with unicode code points.	
out	output zero terminated string that must have room for at least 63 characters plus the terminating zero.	
flags	an <b>Idna_flags</b> value, e.g., <b>IDNA_ALLOW_UNASSIGNED</b> or <b>IDNA_USE_STD3_ASCII_RULES</b> .	

Returns

Returns 0 on success, or an **Idna\_rc** error code.

idna\_to\_unicode\_44i ()

```
int
idna_to_unicode_44i (const uint32_t *in,
                    size_t inlen,
                    uint32_t *out,
                    size_t *outlen,
                    int flags);
```

The ToUnicode operation takes a sequence of Unicode code points that make up one domain label and returns a sequence of Unicode code points. If the input sequence is a label in ACE form, then the result is an equivalent internationalized label that is not in ACE form, otherwise the original sequence is returned unaltered.

ToUnicode never fails. If any step fails, then the original input sequence is returned immediately in that step.

The Punycode decoder can never output more code points than it inputs, but Nameprep can, and therefore ToUnicode can. Note that the number of octets needed to represent a sequence of code points depends on the particular character encoding used.

The inputs to ToUnicode are a sequence of code points, the AllowUnassigned flag, and the UseSTD3ASCIIRules flag. The output of ToUnicode is always a sequence of Unicode code points.

Parameters

in	input array with unicode code points.	
inlen	length of input array with unicode code points.	
out	output array with unicode code points.	



outlen	on input, maximum size of output array with unicode code points, on exit, actual size of output array with unicode code points.	
flags	an <b>Idna_flags</b> value, e.g., <b>IDNA_ALLOW_UNASSIGNED</b> or <b>IDNA_USE_STD3_ASCII_RULES</b> .	

Returns

Returns **Idna\_rc** error condition, but it must only be used for debugging purposes. The output buffer is always guaranteed to contain the correct data according to the specification (sans malloc induced errors). NB! This means that you normally ignore the return code from this function, as checking it means breaking the standard.

idna\_to\_ascii\_4z ()

```
int
idna_to_ascii_4z (const uint32_t *input,
                  char **output,
                  int flags);
```

Convert UCS-4 domain name to ASCII string. The domain name may contain several labels, separated by dots. The output buffer must be deallocated by the caller.

Parameters

input	zero terminated input Unicode string.	
output	pointer to newly allocated output string.	
flags	an <b>Idna_flags</b> value, e.g., <b>IDNA_ALLOW_UNASSIGNED</b> or <b>IDNA_USE_STD3_ASCII_RULES</b> .	

Returns

Returns **IDNA\_SUCCESS** on success, or error code.

idna\_to\_ascii\_8z ()

```
int
idna_to_ascii_8z (const char *input,
                  char **output,
                  int flags);
```

Convert UTF-8 domain name to ASCII string. The domain name may contain several labels, separated by dots. The output buffer must be deallocated by the caller.

Parameters

input	zero terminated input UTF-8 string.	
output	pointer to newly allocated output string.	
flags	an <code>Idna_flags</code> value, e.g., <code>IDNA_ALLOW_UNASSIGNED</code> or <code>IDNA_USE_STD3_ASCII_RULES</code> .	

**Returns**

Returns `IDNA_SUCCESS` on success, or error code.

**idna\_to\_ascii\_lz ()**

```
int
idna_to_ascii_lz (const char *input,
                  char **output,
                  int flags);
```

Convert domain name in the locale's encoding to ASCII string. The domain name may contain several labels, separated by dots. The output buffer must be deallocated by the caller.

**Parameters**

input	zero terminated input string encoded in the current locale's character set.	
output	pointer to newly allocated output string.	
flags	an <code>Idna_flags</code> value, e.g., <code>IDNA_ALLOW_UNASSIGNED</code> or <code>IDNA_USE_STD3_ASCII_RULES</code> .	

**Returns**

Returns `IDNA_SUCCESS` on success, or error code.

**idna\_to\_unicode\_4z4z ()**

```
int
idna_to_unicode_4z4z (const uint32_t *input,
                      uint32_t **output,
                      int flags);
```

Convert possibly ACE encoded domain name in UCS-4 format into a UCS-4 string. The domain name may contain several labels, separated by dots. The output buffer must be deallocated by the caller.

**Parameters**

input	zero-terminated Unicode string.	
output	pointer to newly allocated output Unicode string.	
flags	an <code>Idna_flags</code> value, e.g., <code>IDNA_ALLOW_UNASSIGNED</code> or <code>IDNA_USE_STD3_ASCII_RULES</code> .	

### Returns

Returns `IDNA_SUCCESS` on success, or error code.

### idna\_to\_unicode\_8z4z ()

```
int
idna_to_unicode_8z4z (const char *input,
                     uint32_t **output,
                     int flags);
```

Convert possibly ACE encoded domain name in UTF-8 format into a UCS-4 string. The domain name may contain several labels, separated by dots. The output buffer must be deallocated by the caller.

### Parameters

input	zero-terminated UTF-8 string.	
output	pointer to newly allocated output Unicode string.	
flags	an <code>Idna_flags</code> value, e.g., <code>IDNA_ALLOW_UNASSIGNED</code> or <code>IDNA_USE_STD3_ASCII_RULES</code> .	

### Returns

Returns `IDNA_SUCCESS` on success, or error code.

### idna\_to\_unicode\_8z8z ()

```
int
idna_to_unicode_8z8z (const char *input,
                     char **output,
                     int flags);
```

Convert possibly ACE encoded domain name in UTF-8 format into a UTF-8 string. The domain name may contain several labels, separated by dots. The output buffer must be deallocated by the caller.

### Parameters

input	zero-terminated UTF-8 string.	
-------	-------------------------------	--

output	pointer to newly allocated output UTF-8 string.	
flags	an <code>Idna_flags</code> value, e.g., <code>IDNA_ALLOW_UNASSIGNED</code> or <code>IDNA_USE_STD3_ASCII_RULES</code> .	

### Returns

Returns `IDNA_SUCCESS` on success, or error code.

### idna\_to\_unicode\_8z1z ()

```
int
idna_to_unicode_8z1z (const char *input,
                      char **output,
                      int flags);
```

Convert possibly ACE encoded domain name in UTF-8 format into a string encoded in the current locale's character set. The domain name may contain several labels, separated by dots. The output buffer must be deallocated by the caller.

### Parameters

input	zero-terminated UTF-8 string.	
output	pointer to newly allocated output string encoded in the current locale's character set.	
flags	an <code>Idna_flags</code> value, e.g., <code>IDNA_ALLOW_UNASSIGNED</code> or <code>IDNA_USE_STD3_ASCII_RULES</code> .	

### Returns

Returns `IDNA_SUCCESS` on success, or error code.

### idna\_to\_unicode\_lz1z ()

```
int
idna_to_unicode_lz1z (const char *input,
                      char **output,
                      int flags);
```

Convert possibly ACE encoded domain name in the locale's character set into a string encoded in the current locale's character set. The domain name may contain several labels, separated by dots. The output buffer must be deallocated by the caller.

### Parameters

input	zero-terminated string encoded in the current locale's character set.	
output	pointer to newly allocated output string encoded in the current locale's character set.	
flags	an <code>Idna_flags</code> value, e.g., <code>IDNA_ALLOW_UNASSIGNED</code> or <code>IDNA_USE_STD3_ASCII_RULES</code> .	

Returns

Returns `IDNA_SUCCESS` on success, or error code.

Types and Values

IDNAPI

```
#define IDNAPI
```

enum Idna\_rc

Enumerated return codes of `idna_to_ascii_4i()`, `idna_to_unicode_44i()` functions (and functions derived from those functions). The value 0 is guaranteed to always correspond to success.

Members

IDNA_SUCCESS	Successful operation. This value is guaranteed to always be zero, the remaining ones are only guaranteed to hold non-zero values, for logistical comparison purposes.
IDNA_STRINGPREP_ERROR	Error during string preparation.

IDNA_PUNYCODE_ERROR	Error during punycode operation.
IDNA_CONTAINS_NON_LDH	For IDNA_USE_STD3_ASCII_RULES, indicate that the string contains non-LDH ASCII characters.
IDNA_CONTAINS_LDH	Same as <i>IDNA_CONTAINS_NON_LDH</i> , for compatibility with typographical in earlier versions.

IDNA_CONTAINS_MINUS	For IDNA_USE_STD3_ASCII_RULES, indicate that the string contains a leading or trailing hyphen-minus (U+002D).
IDNA_INVALID_LENGTH	The final output string is not within the (inclusive) range 1 to 63 characters.
IDNA_NO_ACE_PREFIX	The string does not contain the ACE prefix (for ToUnicode).



IDNA_ROUNDTRIP_VERIFY_ERROR	The ToASCII operation on output string does not equal the input.
IDNA_CONTAINS_ACE_PREFIX	The input contains the ACE prefix (for ToASCII).
IDNA_ICONV_ERROR	Character encoding conversion error.
IDNA_MALLOC_ERROR	Could not allocate buffer (this is typically a fatal error).

IDNA\_DLOPEN\_ERROR

Could not dlopen the lib-cidn DSO (only used internally in libc).

enum Idna\_flags

Flags to pass to `idna_to_ascii_4i()`, `idna_to_unicode_44i()` etc.

Members

IDNA\_ALLOW\_UNASSIGNED

Don't reject strings containing unassigned Unicode code points.

IDNA\_USE\_STD3\_ASCII\_RULES

Validate strings according to STD3 rules (i.e., normal host name rules).

IDNA\_ACE\_PREFIX

```
# define IDNA_ACE_PREFIX "xn--"
```

The IANA allocated prefix to use for IDNA. "xn--"

## 1.2 stringprep

stringprep —

### Types and Values

#define	IDNAPI
#define	STRINGPREP_VERSION
enum	Stringprep_rc
enum	Stringprep_profile_flags
enum	Stringprep_profile_steps
#define	STRINGPREP_MAX_MAP_CHARS

### Description

### Functions

### Types and Values

#### IDNAPI

```
#define IDNAPI
```

#### STRINGPREP\_VERSION

```
# define STRINGPREP_VERSION "1.37"
```

String defined via CPP denoting the header file version number. Used together with [stringprep\\_check\\_version\(\)](#) to verify header file and run-time library consistency.

#### enum Stringprep\_rc

Enumerated return codes of [stringprep\(\)](#), [stringprep\\_profile\(\)](#) functions (and macros using those functions). The value 0 is guaranteed to always correspond to success.

#### Members

STRINGPREP\_OK

Successful  
op-  
er-  
a-  
tion.  
This  
value  
is  
guar-  
an-  
teed  
to  
al-  
ways  
be  
zero,  
the  
re-  
main-  
ing  
ones  
are  
only  
guar-  
an-  
teed  
to  
hold  
non-  
zero  
val-  
ues,  
for  
log-  
i-  
cal  
com-  
par-  
i-  
son  
pur-  
poses.

STRINGPREP_CONTAINS_UNASSIGNED	String contain unassigned Unicode code points, which is forbidden by the profile.
STRINGPREP_CONTAINS_PROHIBITED	String contain code points prohibited by the profile.
STRINGPREP_BIDI_BOTH_L_AND_RAL	String contain code points with conflicting bidirectionality category.

STRINGPREP_BIDI_LEADTRAIL_NOT_RAL	Leading and trailing character in string not of proper bidi-rectional category.
STRINGPREP_BIDI_CONTAINS_PROHIBITED	Contains prohibited code points detected by bidi-rectional code.
STRINGPREP_TOO_SMALL_BUFFER	Buffer handed to function was too small. This usually indicate a problem in the calling application.

STRINGPREP\_PROFILE\_ERROR

The stringprep profile was inconsistent. This usually indicate an internal error in the library.

STRINGPREP\_FLAG\_ERROR

The supplied flag conflicted with profile. This usually indicate a problem in the calling application.

STRINGPREP_UNKNOWN_PROFILE	The supplied profile name was not known to the library.
STRINGPREP_ICONV_ERROR	Character encoding conversion error.
STRINGPREP_NFKC_FAILED	The Unicode NFKC operation failed. This usually indicate an internal error in the library.



STRINGPREP\_MALLOC\_ERROR

The  
mal-  
loc()  
was  
out  
of  
mem-  
ory.  
This  
is  
usu-  
ally  
a  
fa-  
tal  
er-  
ror.

**enum Stringprep\_profile\_flags**

Stringprep profile flags.

**Members**

STRINGPREP\_NO\_NFKC

Disable the NFKC normalization, as well as selecting the non-NFKC case folding tables. Usually the profile specifies BIDI and NFKC settings, and applications should not override it unless in special situations.

STRINGPREP\_NO\_BIDI

Disable the BIDI step. Usually the profile specifies BIDI and NFKC settings, and applications should not override it unless in special situations.

STRINGPREP\_NO\_UNASSIGNED

Make the library return with an error if string contains unassigned characters according to profile.

enum Stringprep\_profile\_steps

Various steps in the stringprep algorithm. You really want to study the source code to understand this one. Only useful if you want to add another profile.

Members

STRINGPREP_NFKC	The NFKC step.
STRINGPREP_BIDI	The BIDI step.
STRINGPREP_MAP_TABLE	The MAP step.
STRINGPREP_UNASSIGNED_TABLE	The Unassigned step.
STRINGPREP_PROHIBIT_TABLE	The Prohibited step.
STRINGPREP_BIDI_PROHIBIT_TABLE	The BIDI-Prohibited step.

STRINGPREP_BIDI_RAL_TABLE	The BIDI-RAL step.
STRINGPREP_BIDI_L_TABLE	The BIDI-L step.
STRINGPREP_MAX_MAP_CHARS	

```
# define STRINGPREP_MAX_MAP_CHARS 4
```

Maximum number of code points that can replace a single code point, during stringprep mapping.

### 1.3 punycode

punycode —

#### Functions

const char *	punycode_strerror ()
int	punycode_encode ()
int	punycode_decode ()

#### Types and Values

#define	IDNAPI
enum	Punycode_status
typedef	punycode_uint

#### Description

#### Functions

**punycode\_strerror ()**

```
const char~*  
punycode_strerror (Punycode_status rc);
```

Convert a return code integer to a text string. This string can be used to output a diagnostic message to the user.

PUNYCODE\_SUCCESS: Successful operation. This value is guaranteed to always be zero, the remaining ones are only guaranteed to hold non-zero values, for logical comparison purposes. PUNYCODE\_BAD\_INPUT: Input is invalid. PUNYCODE\_BIG\_OUTPUT: Output would exceed the space provided. PUNYCODE\_OVERFLOW: Input needs wider integers to process.

#### Parameters

rc

an **Punycode\_status** return code.

Returns

Returns a pointer to a statically allocated string containing a description of the error with the return code *rc* .

**punycode\_encode ()**

```
int
punycode_encode (size_t input_length,
                 const punycode_uint input[],
                 const unsigned char case_flags[],
                 size_t *output_length,
                 char output[]);
```

Converts a sequence of code points (presumed to be Unicode code points) to Punycode.

Parameters

input_length	The number of code points in the <i>input</i> array and the number of flags in the <i>case_flags</i> array.
input	An array of code points. They are presumed to be Unicode code points, but that is not strictly REQUIRED. The array contains code points, not code units. UTF-16 uses code units D800 through DFFF to refer to code points 10000..10FFFF. The code points D800..DFFF do not occur in any valid Unicode string. The code points that can occur in Unicode strings (0..D7FF and E000..10FFFF) are also called Unicode scalar values.

case_flags	A <b>NULL</b> pointer or an array of boolean values parallel to the <i>input</i> array. Nonzero (true, flagged) suggests that the corresponding Unicode character be forced to uppercase after being decoded (if possible), and zero (false, unflagged) suggests that it be forced to lowercase (if possible). ASCII code points (0..7F) are encoded literally, except that ASCII letters are forced to uppercase or lowercase according to the corresponding case flags. If <i>case_flags</i> is a <b>NULL</b> pointer then ASCII letters are left as they are, and other code points are treated as unflagged.	
output_length	The caller passes in the maximum number of ASCII code points that it can receive. On successful return it will contain the number of ASCII code points actually output.	
output	An array of ASCII code points. It is <i>*not*</i> null-terminated; it will contain zeros if and only if the <i>input</i> contains zeros. (Of course the caller can leave room for a terminator and add one if needed.)	

## Returns

The return value can be any of the **Punycode\_status** values defined above except **PUNYCODE\_BAD\_INPUT**. If not **PUNYCODE\_SUCCESS**, then *output\_size* and *output* might contain garbage.

## punycode\_decode ()

```
int
punycode_decode (size_t input_length,
                 const char input[],
                 size_t *output_length,
                 punycode_uint output[],
                 unsigned char case_flags[]);
```

Converts Punycode to a sequence of code points (presumed to be Unicode code points).

## Parameters

input_length	The number of ASCII code points in the <i>input</i> array.	
input	An array of ASCII code points (0..7F).	
output_length	The caller passes in the maximum number of code points that it can receive into the <i>output</i> array (which is also the maximum number of flags that it can receive into the <i>case_flags</i> array, if <i>case_flags</i> is not a <b>NULL</b> pointer). On successful return it will contain the number of code points actually output (which is also the number of flags actually output, if <i>case_flags</i> is not a null pointer). The decoder will never need to output more code points than the number of ASCII code points in the input, because of the way the encoding is defined. The number of code points output cannot exceed the maximum possible value of a punycode_uint, even if the supplied <i>output_length</i> is greater than that.	
output	An array of code points like the input argument of <b>punycode_encode()</b> (see above).	
case_flags	A <b>NULL</b> pointer (if the flags are not needed by the caller) or an array of boolean values parallel to the <i>output</i> array. Nonzero (true, flagged) suggests that the corresponding Unicode character be forced to uppercase by the caller (if possible), and zero (false, unflagged) suggests that it be forced to lowercase (if possible). ASCII code points (0..7F) are output already in the proper case, but their flags will be set appropriately so that applying the flags would be harmless.	



## Returns

The return value can be any of the **Punycode\_status** values defined above. If not **PUNYCODE\_SUCCESS**, then *output\_length*, *output*, and *case\_flags* might contain garbage.

## Types and Values

### IDNAPI

```
#define IDNAPI
```

### enum Punycode\_status

Enumerated return codes of **punycode\_encode()** and **punycode\_decode()**. The value 0 is guaranteed to always correspond to success.

### Members

	Successful op- er- a- tion. This value is guar- an- teed to al- ways be zero, the re- main- ing ones are only guar- an- teed to hold non- zero val- ues, for log- i- cal com- par- i- son pur- poses.
PUNYCODE_SUCCESS	
PUNYCODE_BAD_INPUT	Input is in- valid.
PUNYCODE_BIG_OUTPUT	Output would ex- ceed the space pro- vided.

PUNYCODE\_OVERFLOW

Input  
needs  
wider  
in-  
te-  
gers  
to  
pro-  
cess.

punycode\_uint

```
typedef uint32_t punycode_uint;
```

Unicode code point data type, this is always a 32 bit unsigned integer.

1.4 pr29

pr29 —

Functions

const <b>char</b> *	<b>pr29_strerror</b> ()
<b>int</b>	<b>pr29_4</b> ()
<b>int</b>	<b>pr29_4z</b> ()
<b>int</b>	<b>pr29_8z</b> ()

Types and Values

#define	<b>IDNAPI</b>
enum	<b>Pr29_rc</b>

Description

Functions

pr29\_strerror ()

```
const char~*  
pr29_strerror (Pr29_rc rc);
```

Convert a return code integer to a text string. This string can be used to output a diagnostic message to the user.

PR29\_SUCCESS: Successful operation. This value is guaranteed to always be zero, the remaining ones are only guaranteed to hold non-zero values, for logical comparison purposes. PR29\_PROBLEM: A problem sequence was encountered. PR29\_STRINGPREP\_ERROR: The character set conversion failed (only for **pr29\_8z()**).

Parameters

rc	an <b>Pr29_rc</b> return code.	
----	--------------------------------	--

**Returns**

Returns a pointer to a statically allocated string containing a description of the error with the return code *rc*.

**pr29\_4 ()**

```
int
pr29_4 (const uint32_t *in,
        size_t len);
```

Check the input to see if it may be normalized into different strings by different NFKC implementations, due to an anomaly in the NFKC specifications.

**Parameters**

in	input array with unicode code points.	
len	length of input array with unicode code points.	

**Returns**

Returns the **Pr29\_rc** value **PR29\_SUCCESS** on success, and **PR29\_PROBLEM** if the input sequence is a "problem sequence" (i.e., may be normalized into different strings by different implementations).

**pr29\_4z ()**

```
int
pr29_4z (const uint32_t *in);
```

Check the input to see if it may be normalized into different strings by different NFKC implementations, due to an anomaly in the NFKC specifications.

**Parameters**

in	zero terminated array of Unicode code points.	
----	---	--

**Returns**

Returns the **Pr29\_rc** value **PR29\_SUCCESS** on success, and **PR29\_PROBLEM** if the input sequence is a "problem sequence" (i.e., may be normalized into different strings by different implementations).

**pr29\_8z ()**

```
int
pr29_8z (const char *in);
```

Check the input to see if it may be normalized into different strings by different NFKC implementations, due to an anomaly in the NFKC specifications.

---

## Parameters

in	zero terminated input UTF-8 string.
----	--

## Returns

Returns the `Pr29_rc` value `PR29_SUCCESS` on success, and `PR29_PROBLEM` if the input sequence is a "problem sequence" (i.e., may be normalized into different strings by different implementations), or `PR29_STRINGPREP_ERROR` if there was a problem converting the string from UTF-8 to UCS-4.

## Types and Values

### IDNAPI

```
#define IDNAPI
```

### enum Pr29\_rc

Enumerated return codes for `pr29_4()`, `pr29_4z()`, `pr29_8z()`. The value 0 is guaranteed to always correspond to success.

### Members

PR29_SUCCESS	<p>Successful operation. This value is guaranteed to always be zero, the remaining ones are only guaranteed to hold non-zero values, for logical comparison purposes.</p>
PR29_PROBLEM	<p>A problem sequence was encountered.</p>

PR29\_STRINGPREP\_ERROR

The  
char-  
ac-  
ter  
set  
con-  
ver-  
sion  
failed  
(only  
for  
pr29\_8z()).

1.5 tld

tld —

Types and Values

#define IDNAPI

Description

Functions

Types and Values

IDNAPI

```
#define IDNAPI
```

1.6 idn-free

idn-free —

Types and Values

#define IDNAPI

Description

Functions

Types and Values

IDNAPI

```
#define IDNAPI
```



## Chapter 2

# Index

### I

IDNA\_ACE\_PREFIX, [14](#)  
Idna\_flags, [14](#)  
Idna\_rc, [9](#)  
idna\_strerror, [3](#)  
idna\_to\_ascii\_4i, [3](#)  
idna\_to\_ascii\_4z, [5](#)  
idna\_to\_ascii\_8z, [5](#)  
idna\_to\_ascii\_lz, [6](#)  
idna\_to\_unicode\_44i, [4](#)  
idna\_to\_unicode\_4z4z, [6](#)  
idna\_to\_unicode\_8z4z, [7](#)  
idna\_to\_unicode\_8z8z, [7](#)  
idna\_to\_unicode\_8zlz, [8](#)  
idna\_to\_unicode\_lzlz, [8](#)  
IDNAPI, [9](#), [15](#), [29](#), [33](#), [35](#)

### P

pr29\_4, [32](#)  
pr29\_4z, [32](#)  
pr29\_8z, [32](#)  
Pr29\_rc, [33](#)  
pr29\_strerror, [31](#)  
punycode\_decode, [27](#)  
punycode\_encode, [26](#)  
Punycode\_status, [29](#)  
punycode\_strerror, [25](#)  
punycode\_uint, [31](#)

### S

STRINGPREP\_MAX\_MAP\_CHARS, [25](#)  
Stringprep\_profile\_flags, [21](#)  
Stringprep\_profile\_steps, [24](#)  
Stringprep\_rc, [15](#)  
STRINGPREP\_VERSION, [15](#)

---