



ASN.1

Copyright © 1997-2016 Ericsson AB. All Rights Reserved.
ASN.1 3.0.4
April 4, 2016

Copyright © 1997-2016 Ericsson AB. All Rights Reserved.

The contents of this file are subject to the Erlang Public License, Version 1.1, (the "License"); you may not use this file except in compliance with the License. You should have received a copy of the Erlang Public License along with this software. If not, it can be retrieved online at <http://www.erlang.org/>. Software distributed under the License is distributed on an "AS IS" basis, WITHOUT WARRANTY OF ANY KIND, either express or implied. See the License for the specific language governing rights and limitations under the License. Ericsson AB. All Rights Reserved..

April 4, 2016



1 Asn1 User's Guide

The *Asn1* application contains modules with compile-time and run-time support for ASN.1.

1.1 Asn1

1.1.1 Introduction

Features

The *Asn1* application provides:

- An ASN.1 compiler for Erlang, which generates encode and decode functions to be used by Erlang programs sending and receiving ASN.1 specified data.
- Run-time functions used by the generated code.
- Support for the following encoding rules:
 - Basic Encoding Rules (*BER*)
 - Distinguished Encoding Rules (*DER*), a specialized form of BER that is used in security-conscious applications.
 - Packed Encoding Rules (*PER*); both the aligned and unaligned variant.

Overview

ASN.1 (Abstract Syntax Notation One) is a formal language for describing data structures to be exchanged between distributed computer systems. The purpose of ASN.1 is to have a platform and programming language independent notation to express types using a standardized set of rules for the transformation of values of a defined type into a stream of bytes. This stream of bytes can then be sent on any type of communication channel. This way, two applications written in different programming languages running on different computers with different internal representation of data can exchange instances of structured data types.

Prerequisites

It is assumed that the reader is familiar with the ASN.1 notation as documented in the standard definition [1] which is the primary text. It may also be helpful, but not necessary, to read the standard definitions [2] [3] [4] [5].

A good book explaining those reference texts is [6], which is free to download at <http://www.oss.com/asn1/dubuisson.html>.

Capabilities

This application covers all features of ASN.1 up to the 1997 edition of the specification. In the 2002 edition of ASN.1 a number of new features were introduced. The following features of the 2002 edition are fully or partly supported as shown below:

- Decimal notation (e.g., "1.5e3") for REAL values. The NR1, NR2 and NR3 formats as explained in ISO6093 are supported.
- The RELATIVE-OID type for relative object identifiers is fully supported.
- The subtype constraint (CONTAINING/ENCODED BY) to constrain the content of an octet string or a bit string is parsed when compiling, but no further action is taken. This constraint is not a PER-visible constraint.

- The subtype constraint by regular expressions (PATTERN) for character string types is parsed when compiling, but no further action is taken. This constraint is not a PER-visible constraint.
- Multiple-line comments as in C, `/* ... */`, are supported.

1.1.2 Getting Started with Asn1

A First Example

The following example demonstrates the basic functionality used to run the Erlang ASN.1 compiler.

Create a file called `People.asn` containing the following:

```
People DEFINITIONS AUTOMATIC TAGS ::=
BEGIN
  Person ::= SEQUENCE {
    name PrintableString,
    location INTEGER {home(0),field(1),roving(2)},
    age INTEGER OPTIONAL
  }
END
```

This file (`People.asn`) must be compiled before it can be used. The ASN.1 compiler checks that the syntax is correct and that the text represents proper ASN.1 code before generating an abstract syntax tree. The code-generator then uses the abstract syntax tree in order to generate code.

The generated Erlang files will be placed in the current directory or in the directory specified with the `{outdir,Dir}` option. The following shows how the compiler can be called from the Erlang shell:

```
1> asn1ct:compile("People", [ber]).
ok
2>
```

The verbose option can be given to have information about the generated files printed:

```
2> asn1ct:compile("People", [ber,verbose]).
Erlang ASN.1 compiling "People.asn"
--{generated,"People.asnldb"}--
--{generated,"People.hrl"}--
--{generated,"People.erl"}--
ok
3>
```

The ASN.1 module `People` is now accepted and the abstract syntax tree is saved in the `People.asnldb` file; the generated Erlang code is compiled using the Erlang compiler and loaded into the Erlang run-time system. Now there is an API for `encode/2` and `decode/2` in the module `People`, which is invoked by:

```
'People':encode(<Type name>, <Value>)
or
'People':decode(<Type name>, <Value>)
```

Assume there is a network application which receives instances of the ASN.1 defined type `Person`, modifies and sends them back again:

1.1 Asn1

```
receive
  {Port,{data,Bytes}} ->
    case 'People':decode('Person',Bytes) of
      {ok,P} ->
        {ok,Answer} = 'People':encode('Person',mk_answer(P)),
        Port ! {self(),{command,Answer}};
      {error,Reason} ->
        exit({error,Reason})
    end
end,
```

In the example above, a series of bytes is received from an external source and the bytes are then decoded into a valid Erlang term. This was achieved with the call `'People':decode('Person',Bytes)` which returned an Erlang value of the ASN.1 type `Person`. Then an answer was constructed and encoded using `'People':encode('Person',Answer)` which takes an instance of a defined ASN.1 type and transforms it to a binary according to the BER or PER encoding rules.

The encoder and the decoder can also be run from the shell.

```
2> Rockstar = {'Person',"Some Name",roving,50}.
{'Person',"Some Name",roving,50}
3> {ok,Bin} = 'People':encode('Person',Rockstar).
{ok,<<243,17,19,9,83,111,109,101,32,78,97,109,101,2,1,2,
    2,1,50>>}
4> {ok,Person} = 'People':decode('Person',Bin).
{ok,{'Person',"Some Name",roving,50}}
5>
```

Module dependencies

It is common that ASN.1 modules import defined types, values and other entities from another ASN.1 module.

Earlier versions of the ASN.1 compiler required that modules that were imported from had to be compiled before the module that imported. This caused problems when ASN.1 modules had circular dependencies.

Referenced modules are now parsed when the compiler finds an entity that is imported. There will not be any code generated for the referenced module. However, the compiled module rely on that the referenced modules also will be compiled.

1.1.3 The Asn1 Application User Interface

The Asn1 application provides two separate user interfaces:

- The module `asn1ct` which provides the compile-time functions (including the compiler).
- The module `asn1rt_nif` which provides the run-time functions for the ASN.1 decoder for the BER back-end.

The reason for the division of the interface into compile-time and run-time is that only run-time modules (`asn1rt*`) need to be loaded in an embedded system.

Compile-time Functions

The ASN.1 compiler can be invoked directly from the command-line by means of the `erlc` program. This is convenient when compiling many ASN.1 files from the command-line or when using Makefiles. Here are some examples of how the `erlc` command can be used to invoke the ASN.1 compiler:

```
erlc Person.asn
erlc -bper Person.asn
erlc -bber ../Example.asn
```

```
erlc -o ../asnfiles -I ../asnfiles -I /usr/local/standards/asn1 Person.asn
```

The useful options for the ASN.1 compiler are:

`-b[ber | per | uper]`

Choice of encoding rules, if omitted `ber` is the default.

`-o OutDirectory`

Where to put the generated files, default is the current directory.

`-I IncludeDir`

Where to search for `.asn1db` files and ASN.1 source specs in order to resolve references to other modules. This option can be repeated many times if there are several places to search in. The compiler will always search the current directory first.

`+der`

DER encoding rule. Only when using `-ber` option.

`+asn1config`

This functionality works together with the `ber` option. It enables the specialized decodes, see the *Specialized Decode* chapter.

`+undec_rest`

A buffer that holds a message being decoded may also have trailing bytes. If those trailing bytes are important they can be returned along with the decoded value by compiling the ASN.1 specification with the `+undec_rest` option. The return value from the decoder will be `{ok, Value, Rest}` where `Rest` is a binary containing the trailing bytes.

`+'Any Erlc Option'`

You may add any option to the Erlang compiler when compiling the generated Erlang files. Any option unrecognized by the ASN.1 compiler will be passed to the Erlang compiler.

For a complete description of `erlc` see *Erts Reference Manual*.

The compiler and other compile-time functions can also be invoked from the Erlang shell. Below follows a brief description of the primary functions, for a complete description of each function see *the Asn1 Reference Manual*, the `asn1ct` module.

The compiler is invoked by using `asn1ct:compile/1` with default options, or `asn1ct:compile/2` if explicit options are given. Example:

```
asn1ct:compile("H323-MESSAGES.asn1").
```

which equals:

```
asn1ct:compile("H323-MESSAGES.asn1",[ber]).
```

If one wants PER encoding:

```
asn1ct:compile("H323-MESSAGES.asn1",[per]).
```

1.1 Asn1

The generic encode and decode functions can be invoked like this:

```
'H323-MESSAGES':encode('SomeChoiceType',{call,"octetstring"}).  
'H323-MESSAGES':decode('SomeChoiceType',Bytes).
```

Run-time Functions

When an ASN.1 specification is compiled with the `ber` option, the module `asn1rt_nif` module and the NIF library in `asn1/priv_dir` will be needed at run-time.

By invoking the function `info/0` in a generated module, one gets information about which compiler options were used.

Errors

Errors detected at compile time appear on the screen together with a line number indicating where in the source file the error was detected. If no errors are found, an Erlang ASN.1 module will be created.

The run-time encoders and decoders execute within a catch and returns `{ok, Data}` or `{error, {asn1, Description}}` where `Description` is an Erlang term describing the error.

1.1.4 Multi-file Compilation

There are various reasons for using multi-file compilation:

- You want to choose the name for the generated module, perhaps because you need to compile the same specs for different encoding rules.
- You want only one resulting module.

You need to specify which ASN.1 specs you will compile in a module that must have the extension `.set.asn`. You chose name of the module and provide the names of the ASN.1 specs. For instance, if you have the specs `File1.asn`, `File2.asn` and `File3.asn` your module `MyModule.set.asn` will look like:

```
File1.asn  
File2.asn  
File3.asn
```

If you compile with:

```
~> erlc MyModule.set.asn
```

the result will be one merged module `MyModule.erl` with the generated code from the three ASN.1 specs.

1.1.5 A quick note about tags

Tags used to be important for all users of ASN.1, because it was necessary to manually add tags to certain constructs in order for the ASN.1 specification to be valid. Here is an example of an old-style specification:

```
Tags DEFINITIONS ::=  
BEGIN  
  Afters ::= CHOICE { cheese [0] IA5String,  
                      dessert [1] IA5String }
```



```
END
```

Without the tags (the numbers in square brackets) the ASN.1 compiler would refuse to compile the file.

In 1994 the global tagging mode AUTOMATIC TAGS was introduced. By putting AUTOMATIC TAGS in the module header, the ASN.1 compiler will automatically add tags when needed. Here is the same specification in AUTOMATIC TAGS mode:

```
Tags DEFINITIONS AUTOMATIC TAGS ::=
BEGIN
  Afters ::= CHOICE { cheese IA5String,
                      dessert IA5String }
END
```

Tags will not be mentioned any more in this manual.

1.1.6 The ASN.1 Types

This section describes the ASN.1 types including their functionality, purpose and how values are assigned in Erlang. ASN.1 has both primitive and constructed types:

<i>Primitive types</i>	<i>Constructed types</i>
<i>BOOLEAN</i>	<i>SEQUENCE</i>
<i>INTEGER</i>	<i>SET</i>
<i>REAL</i>	<i>CHOICE</i>
<i>NULL</i>	<i>SET OF and SEQUENCE OF</i>
<i>ENUMERATED</i>	<i>ANY</i>
<i>BIT STRING</i>	<i>ANY DEFINED BY</i>
<i>OCTET STRING</i>	<i>EXTERNAL</i>
<i>Character Strings</i>	<i>EMBEDDED PDV</i>
<i>OBJECT IDENTIFIER</i>	<i>CHARACTER STRING</i>
<i>Object Descriptor</i>	
<i>The TIME types</i>	

Table 1.1: The supported ASN.1 types

1.1 Asn1

Note:

Values of each ASN.1 type has its own representation in Erlang described in the following subsections. Users shall provide these values for encoding according to the representation, as in the example below.

```
Operational ::= BOOLEAN --ASN.1 definition
```

In Erlang code it may look like:

```
Val = true,  
{ok,Bytes} = MyModule:encode('Operational', Val),
```

Below follows a description of how values of each type can be represented in Erlang.

BOOLEAN

Booleans in ASN.1 express values that can be either TRUE or FALSE. The meanings assigned to TRUE or FALSE is beyond the scope of this text.

In ASN.1 it is possible to have:

```
Operational ::= BOOLEAN
```

Assigning a value to the type Operational in Erlang is possible by using the following Erlang code:

```
Myvar1 = true,
```

Thus, in Erlang the atoms `true` and `false` are used to encode a boolean value.

INTEGER

ASN.1 itself specifies indefinitely large integers, and the Erlang systems with versions 4.3 and higher, support very large integers, in practice indefinitely large integers.

The concept of sub-typing can be applied to integers as well as to other ASN.1 types. The details of sub-typing are not explained here, for further info see []. A variety of syntaxes are allowed when defining a type as an integer:

```
T1 ::= INTEGER  
T2 ::= INTEGER (-2..7)  
T3 ::= INTEGER (0..MAX)  
T4 ::= INTEGER (0<..MAX)  
T5 ::= INTEGER (MIN<..-99)  
T6 ::= INTEGER {red(0),blue(1),white(2)}
```

The Erlang representation of an ASN.1 INTEGER is an integer or an atom if a so called `Named Number List` (see T6 above) is specified.

Below is an example of Erlang code which assigns values for the above types:

```
T1value = 0,
T2value = 6,
T6value1 = blue,
T6value2 = 0,
T6value3 = white
```

The Erlang variables above are now bound to valid instances of ASN.1 defined types. This style of value can be passed directly to the encoder for transformation into a series of bytes.

The decoder will return an atom if the value corresponds to a symbol in the Named Number List.

REAL

The following ASN.1 type is used for real numbers:

```
R1 ::= REAL
```

It can be assigned a value in Erlang as:

```
R1value1 = "2.14",
R1value2 = {256,10,-2},
```

In the last line note that the tuple {256,10,-2} is the real number 2.56 in a special notation, which will encode faster than simply stating the number as "2.56". The arity three tuple is {Mantissa, Base, Exponent} i.e. Mantissa * Base^Exponent.

NULL

Null is suitable in cases where supply and recognition of a value is important but the actual value is not.

```
Notype ::= NULL
```

The NULL type can be assigned in Erlang:

```
N1 = 'NULL',
```

The actual value is the quoted atom 'NULL'.

ENUMERATED

The enumerated type can be used, when the value we wish to describe, may only take one of a set of predefined values.

```
DaysOfTheWeek ::= ENUMERATED {
```

1.1 Asn1

```
sunday(1),monday(2),tuesday(3),  
wednesday(4),thursday(5),friday(6),saturday(7) }
```

For example to assign a weekday value in Erlang use the same atom as in the `Enumerations` of the type definition:

```
Day1 = saturday,
```

The enumerated type is very similar to an integer type, when defined with a set of predefined values. An enumerated type differs from an integer in that it may only have specified values, whereas an integer can also have any other value.

BIT STRING

The BIT STRING type can be used to model information which is made up of arbitrary length series of bits. It is intended to be used for a selection of flags, not for binary files.

In ASN.1 BIT STRING definitions may look like:

```
Bits1 ::= BIT STRING  
Bits2 ::= BIT STRING {foo(0),bar(1),gnu(2),gnome(3),punk(14)}
```

There are two notations available for representation of BIT STRING values in Erlang and as input to the encode functions.

- A bitstring. By default, a BIT STRING with no symbolic names will be decoded to an Erlang bitstring.
- A list of atoms corresponding to atoms in the `NamedBitList` in the BIT STRING definition. A BIT STRING with symbolic names will always be decoded to this format.

Example:

```
Bits1Val1 = <<0:1,1:1,0:1,1:1,1:1>>,  
Bits2Val1 = [gnu,punk],  
Bits2Val2 = <<2#1110:4>>,  
Bits2Val3 = [bar,gnu,gnome],
```

`Bits2Val2` and `Bits2Val3` above denote the same value.

`Bits2Val1` is assigned symbolic values. The assignment means that the bits corresponding to `gnu` and `punk` i.e. bits 2 and 14 are set to 1 and the rest set to 0. The symbolic values appear as a list of values. If a named value appears, which is not specified in the type definition, a run-time error will occur.

BIT STRINGS may also be sub-typed with, for example, a SIZE specification:

```
Bits3 ::= BIT STRING (SIZE(0..31))
```

This means that no bit higher than 31 can ever be set.

Deprecated representations for BIT STRING

In addition to the representations described above, the following deprecated representations are available if the specification has been compiled with the `legacy_erlang_types` option:

- A list of binary digits (0 or 1). This format is accepted as input to the encode functions, and a BIT STRING will be decoded to this format if the *legacy_bit_string* option has been given.
- As {Unused, Binary} where Unused denotes how many trailing zero-bits 0 to 7 that are unused in the least significant byte in Binary. This format is accepted as input to the encode functions, and a BIT STRING will be decoded to this format if *compact_bit_string* has been given.
- A hexadecimal number (or an integer). This format should be avoided, since it is easy to misinterpret a BIT STRING value in this format.

OCTET STRING

The OCTET STRING is the simplest of all ASN.1 types. The OCTET STRING only moves or transfers e.g. binary files or other unstructured information complying to two rules. Firstly, the bytes consist of octets and secondly, encoding is not required.

It is possible to have the following ASN.1 type definitions:

```
O1 ::= OCTET STRING
O2 ::= OCTET STRING (SIZE(28))
```

With the following example assignments in Erlang:

```
O1Val = <<17,13,19,20,0,0,255,254>>,
O2Val = <<"must be exactly 28 chars....">>,
```

By default, an OCTET STRING is always represented as an Erlang binary. If the specification has been compiled with the *legacy_erlang_types* option, the encode functions will accept both lists and binaries, and the decode functions will decode an OCTET STRING to a list.

Character Strings

ASN.1 supports a wide variety of character sets. The main difference between OCTET STRINGS and the Character strings is that OCTET STRINGS have no imposed semantics on the bytes delivered.

However, when using for instance the IA5String (which closely resembles ASCII) the byte 65 (in decimal notation) *means* the character 'A'.

For example, if a defined type is to be a VideotexString and an octet is received with the unsigned integer value X, then the octet should be interpreted as specified in the standard ITU-T T.100,T.101.

The ASN.1 to Erlang compiler will not determine the correct interpretation of each BER (Basic Encoding Rules) string octet value with different Character strings. Interpretation of octets is the responsibility of the application. Therefore, from the BER string point of view, octets appear to be very similar to character strings and are compiled in the same way.

It should be noted that when PER (Packed Encoding Rules) is used, there is a significant difference in the encoding scheme between OCTET STRINGS and other strings. The constraints specified for a type are especially important for PER, where they affect the encoding.

Here are some examples:

```
Digs ::= NumericString (SIZE(1..3))
TextFile ::= IA5String (SIZE(0..64000))
```

1.1 Asn1

with corresponding Erlang assignments:

```
DigsVal1 = "456",
DigsVal2 = "123",
TextFileVal1 = "abc...xyz...",
TextFileVal2 = [88,76,55,44,99,121 ..... a lot of characters here ....]
```

The Erlang representation for "BMPString" and "UniversalString" is either a list of ASCII values or a list of quadruples. The quadruple representation associates to the Unicode standard representation of characters. The ASCII characters are all represented by quadruples beginning with three zeros like {0,0,0,65} for the 'A' character. When decoding a value for these strings the result is a list of quadruples, or integers when the value is an ASCII character.

The following example shows how it works. We have the following specification in the file `PrimStrings.asn1`.

```
PrimStrings DEFINITIONS AUTOMATIC TAGS ::=
BEGIN
    BMP ::= BMPString
END
```

Encoding and decoding some strings:

```
1> asn1ct:compile('PrimStrings', [ber]).
ok
2> {ok,Bytes1} = 'PrimStrings':encode('BMP', [{0,0,53,53},{0,0,45,56}]).
{ok,<<30,4,53,54,45,56>>}
3> 'PrimStrings':decode('BMP', Bytes1).
{ok,[{0,0,53,53},{0,0,45,56}]}
4> {ok,Bytes2} = 'PrimStrings':encode('BMP', [{0,0,53,53},{0,0,0,65}]).
{ok,<<30,4,53,53,0,65>>}
5> 'PrimStrings':decode('BMP', Bytes2).
{ok,[{0,0,53,53},65]}
6> {ok,Bytes3} = 'PrimStrings':encode('BMP', "BMP string").
{ok,<<30,20,0,66,0,77,0,80,0,32,0,115,0,116,0,114,0,105,0,110,0,103>>}
7> 'PrimStrings':decode('BMP', Bytes3).
{ok,"BMP string"}
```

The UTF8String type is represented as a UTF-8 encoded binary in Erlang. Such binaries can be created directly using the binary syntax or by converting from a list of Unicode code points using the `unicode:characters_to_binary/1` function.

Here are some examples showing how UTF-8 encoded binaries can be created and manipulated:

```
1> Gs = "### #####".
[1052,1086,1081,32,1084,1072,1083,1077,1085,1100,1082,1080,
 1081,32,1043,1085,1086,1084]
2> Gbin = unicode:characters_to_binary(Gs).
<<208,156,208,190,208,185,32,208,188,208,176,208,187,208,
 181,208,189,209,140,208,186,208,184,208,185,32,208,147,
 208,...>>
3> Gbin = <<"### #####"/utf8>>.
<<208,156,208,190,208,185,32,208,188,208,176,208,187,208,
 181,208,189,209,140,208,186,208,184,208,185,32,208,147,
 208,...>>
4> Gs = unicode:characters_to_list(Gbin).
```

```
[1052,1086,1081,32,1084,1072,1083,1077,1085,1100,1082,1080,
1081,32,1043,1085,1086,1084]
```

See the *unicode* module for more details.

In the following example we will use this ASN.1 specification:

```
UTF DEFINITIONS AUTOMATIC TAGS ::=
BEGIN
    UTF ::= UTF8String
END
```

Encoding and decoding a string with Unicode characters:

```
5> asn1ct:compile('UTF', [ber]).
ok
6> {ok,Bytes1} = 'UTF':encode('UTF', <<"####"/utf8>>).
{ok,<<12,8,208,147,208,189,208,190,208,188>>}
7> {ok,Bin1} = 'UTF':decode('UTF', Bytes1).
{ok,<<208,147,208,189,208,190,208,188>>}
8> io:format("~ts\n", [Bin1]).
####
ok
9> unicode:characters_to_list(Bin1).
[1043,1085,1086,1084]
```

OBJECT IDENTIFIER

The OBJECT IDENTIFIER is used whenever a unique identity is required. An ASN.1 module, a transfer syntax, etc. is identified with an OBJECT IDENTIFIER. Assume the example below:

```
Oid ::= OBJECT IDENTIFIER
```

Therefore, the example below is a valid Erlang instance of the type 'Oid'.

```
OidVal1 = {1,2,55},
```

The OBJECT IDENTIFIER value is simply a tuple with the consecutive values which must be integers.

The first value is limited to the values 0, 1 or 2 and the second value must be in the range 0..39 when the first value is 0 or 1.

The OBJECT IDENTIFIER is a very important type and it is widely used within different standards to uniquely identify various objects. In [], there is an easy-to-understand description of the usage of OBJECT IDENTIFIER.

Object Descriptor

Values of this type can be assigned a value as an ordinary string like this:

```
"This is the value of an Object descriptor"
```

The TIME Types

Two different time types are defined within ASN.1, Generalized Time and UTC (Universal Time Coordinated), both are assigned a value as an ordinary string within double quotes i.e. "19820102070533.8".

In case of DER encoding the compiler does not check the validity of the time values. The DER requirements upon those strings is regarded as a matter for the application to fulfill.

SEQUENCE

The structured types of ASN.1 are constructed from other types in a manner similar to the concepts of array and struct in C.

A SEQUENCE in ASN.1 is comparable with a struct in C and a record in Erlang. A SEQUENCE may be defined as:

```
Pdu ::= SEQUENCE {  
    a INTEGER,  
    b REAL,  
    c OBJECT IDENTIFIER,  
    d NULL }
```

This is a 4-component structure called 'Pdu'. The major format for representation of SEQUENCE in Erlang is the record format. For each SEQUENCE and SET in an ASN.1 module an Erlang record declaration is generated. For Pdu above, a record like this is defined:

```
-record('Pdu',{a, b, c, d}).
```

The record declarations for a module M are placed in a separate M.hrl file.

Values can be assigned in Erlang as shown below:

```
MyPdu = #'Pdu'{a=22,b=77.99,c={0,1,2,3,4},d='NULL'}.
```

The decode functions will return a record as result when decoding a SEQUENCE or a SET.

A SEQUENCE and a SET may contain a component with a DEFAULT key word followed by the actual value that is the default value. The DEFAULT keyword means that the application doing the encoding can omit encoding of the value, thus resulting in fewer bytes to send to the receiving application.

An application can use the atom `asn1_DEFAULT` to indicate that the encoding should be omitted for that position in the SEQUENCE.

Depending on the encoding rules, the encoder may also compare the given value to the default value and automatically omit the encoding if they are equal. How much effort the encoder makes to compare the values depends on the encoding rules. The DER encoding rules forbids encoding a value equal to the default value, so it has a more thorough and time-consuming comparison than the encoders for the other encoding rules.

In the following example we will use this ASN.1 specification:

```
File DEFINITIONS AUTOMATIC TAGS ::=  
BEGIN
```



```

Seq1 ::= SEQUENCE {
    a INTEGER DEFAULT 1,
    b Seq2 DEFAULT {aa TRUE, bb 15}
}

Seq2 ::= SEQUENCE {
    aa BOOLEAN,
    bb INTEGER
}

Seq3 ::= SEQUENCE {
    bs BIT STRING {a(0), b(1), c(2)} DEFAULT {a, c}
}
END

```

Here is an example where the BER encoder is able to omit encoding of the default values:

```

1> asn1ct:compile('File', [ber]).
ok
2> 'File':encode('Seq1', {'Seq1',asn1_DEFAULT,asn1_DEFAULT}).
{ok,<<48,0>>}
3> 'File':encode('Seq1', {'Seq1',1,{'Seq2',true,15}}).
{ok,<<48,0>>}

```

And here is an example with a named BIT STRING where the BER encoder will not omit the encoding:

```

4> 'File':encode('Seq3', {'Seq3',asn1_DEFAULT}).
{ok,<<48,0>>}
5> 'File':encode('Seq3', {'Seq3',<<16#101:3>>}).
{ok,<<48,4,128,2,5,160>>}

```

The DER encoder will omit the encoding for the same BIT STRING:

```

6> asn1ct:compile('File', [ber,der]).
ok
7> 'File':encode('Seq3', {'Seq3',asn1_DEFAULT}).
{ok,<<48,0>>}
8> 'File':encode('Seq3', {'Seq3',<<16#101:3>>}).
{ok,<<48,0>>}

```

SET

In Erlang, the SET type is used exactly as SEQUENCE. Note that if the BER or DER encoding rules are used, decoding a SET is slower than decoding a SEQUENCE because the components must be sorted.

Notes about extensibility for SEQUENCE and SET

When a SEQUENCE or SET contains an extension marker and extension components like this:

```

SExt ::= SEQUENCE {
    a INTEGER,
    ...,
    b BOOLEAN }

```

1.1 Asn1

It means that the type may get more components in newer versions of the ASN.1 spec. In this case it has got a new component `b`. Thus, incoming messages that will be decoded may have more or fewer components than this one.

The component `b` will be treated as an original component when encoding a message. In this case, as it is not an optional element, it must be encoded.

During decoding the `b` field of the record will get the decoded value of the `b` component if present and otherwise the value `asn1_NOVALUE`.

CHOICE

The CHOICE type is a space saver and is similar to the concept of a 'union' in the C language.

Assume:

```
SomeModuleName DEFINITIONS AUTOMATIC TAGS ::=
BEGIN
T ::= CHOICE {
    x REAL,
    y INTEGER,
    z OBJECT IDENTIFIER }
END
```

It is then possible to assign values:

```
TVal1 = {y,17},
TVal2 = {z,{0,1,2}},
```

A CHOICE value is always represented as the tuple `{ChoiceAlternative, Val}` where `ChoiceAlternative` is an atom denoting the selected choice alternative.

Extensible CHOICE

When a CHOICE contains an extension marker and the decoder detects an unknown alternative of the CHOICE the value is represented as:

```
{asn1_ExtAlt, BytesForOpenType}
```

Where `BytesForOpenType` is a list of bytes constituting the encoding of the "unknown" CHOICE alternative.

SET OF and SEQUENCE OF

The SET OF and SEQUENCE OF types correspond to the concept of an array found in several programming languages. The Erlang syntax for both of these types is straight forward. For example:

```
Arr1 ::= SET SIZE (5) OF INTEGER (4..9)
Arr2 ::= SEQUENCE OF OCTET STRING
```

We may have the following in Erlang:

```
Arr1Val = [4,5,6,7,8],
```

```
Arr2Val = ["abc",[14,34,54],"Octets"],
```

Please note that the definition of the SET OF type implies that the order of the components is undefined, but in practice there is no difference between SET OF and SEQUENCE OF. The ASN.1 compiler for Erlang does not randomize the order of the SET OF components before encoding.

However, in case of a value of the type SET OF, the DER encoding format requires the elements to be sent in ascending order of their encoding, which implies an expensive sorting procedure in run-time. Therefore it is strongly recommended to use SEQUENCE OF instead of SET OF if it is possible.

ANY and ANY DEFINED BY

The types ANY and ANY DEFINED BY have been removed from the standard since 1994. It is recommended not to use these types any more. They may, however, exist in some old ASN.1 modules. The idea with this type was to leave a "hole" in a definition where one could put unspecified data of any kind, even non ASN.1 data.

A value of this type is encoded as an open type.

Instead of ANY/ANY DEFINED BY one should use information object class, table constraints and parameterization. In particular the construct TYPE-IDENTIFIER.@Type accomplish the same as the deprecated ANY.

See also *Information object*

EXTERNAL, EMBEDDED PDV and CHARACTER STRING

These types are used in presentation layer negotiation. They are encoded according to their associated type, see [].

The EXTERNAL type had a slightly different associated type before 1994. [] states that encoding shall follow the older associate type. Therefore does generated encode/decode functions convert values of the newer format to the older format before encoding. This implies that it is allowed to use EXTERNAL type values of either format for encoding. Decoded values are always returned on the newer format.

Embedded Named Types

The structured types previously described may very well have other named types as their components. The general syntax to assign a value to the component C of a named ASN.1 type T in Erlang is the record syntax #'T'{'C'=Value}. Where Value may be a value of yet another type T2.

For example:

```
EmbeddedExample DEFINITIONS AUTOMATIC TAGS ::=
BEGIN
B ::= SEQUENCE {
    a Arr1,
    b T }

Arr1 ::= SET SIZE (5) OF INTEGER (4..9)

T ::= CHOICE {
    x REAL,
    y INTEGER,
    z OBJECT IDENTIFIER }
END
```

The SEQUENCE b can be encoded like this in Erlang:

```
1> 'EmbeddedExample':encode('B', {'B',[4,5,6,7,8],{x,"7.77"}}).
```

```
{ok,<<5,56,0,8,3,55,55,55,46,69,45,50>>}
```

1.1.7 Naming of Records in .hrl Files

When an ASN.1 specification is compiled all defined types of type SET or SEQUENCE will result in a corresponding record in the generated hrl file. This is because the values for SET/SEQUENCE as mentioned in sections above are represented as records.

Though there are some special cases of this functionality that are presented below.

Embedded Structured Types

It is also possible in ASN.1 to have components that are themselves structured types. For example, it is possible to have:

```
Emb ::= SEQUENCE {  
    a SEQUENCE OF OCTET STRING,  
    b SET {  
        a INTEGER,  
        b INTEGER DEFAULT 66},  
    c CHOICE {  
        a INTEGER,  
        b FooType } }  
  
FooType ::= [3] VisibleString
```

The following records are generated because of the type Emb:

```
-record('Emb',{a, b, c}).  
-record('Emb_b',{a, b = asn1_DEFAULT}). % the embedded SET type
```

Values of the Emb type can be assigned like this:

```
V = #'Emb'{a=["qqqq",[1,2,255]],  
          b = #'Emb_b'{a=99},  
          c = {b,"Can you see this"}}.
```

For an embedded type of type SEQUENCE/SET in a SEQUENCE/SET the record name is extended with an underscore and the component name. If the embedded structure is deeper with SEQUENCE, SET or CHOICE types in the line, each component-/alternative-name will be added to the record-name.

For example:

```
Seq ::= SEQUENCE{  
    a CHOICE{  
        b SEQUENCE {  
            c INTEGER  
        }  
    }  
}
```

will result in the following record:

```
-record('Seq_a_b',{c}).
```

If the structured type has a component with an embedded SEQUENCE OF/SET OF which embedded type in turn is a SEQUENCE/SET it will give a record with the SEQOF/SETOF addition as in the following example:

```
Seq ::= SEQUENCE {
    a SEQUENCE OF SEQUENCE {
        b
    }
    c SET OF SEQUENCE {
        d
    }
}
```

This results in the records:

```
-record('Seq_a_SEQOF',{b}).
-record('Seq_c_SETOF',{d}).
```

A parameterized type should be considered as an embedded type. Each time a such type is referenced an instance of it is defined. Thus in the following example a record with name 'Seq_b' is generated in the .hrl file and used to hold values.

```
Seq ::= SEQUENCE {
    b PType{INTEGER}
}

PType{T} ::= SEQUENCE{
    id T
}
```

Recursive Types

Types may refer to themselves. Suppose:

```
Rec ::= CHOICE {
    nothing NULL,
    something SEQUENCE {
        a INTEGER,
        b OCTET STRING,
        c Rec }}
```

This type is recursive; that is, it refers to itself. This is allowed in ASN.1 and the ASN.1-to-Erlang compiler supports this recursive type. A value for this type is assigned in Erlang as shown below:

```
V = {something, #'Rec_something' {a = 77,
                                b = "some octets here",
                                c = {nothing, 'NULL'}}}.
```

1.1.8 ASN.1 Values

Values can be assigned to ASN.1 type within the ASN.1 code itself, as opposed to the actions taken in the previous chapter where a value was assigned to an ASN.1 type in Erlang. The full value syntax of ASN.1 is supported and [X.680] describes in detail how to assign values in ASN.1. Below is a short example:

```
TT ::= SEQUENCE {
    a INTEGER,
    b SET OF OCTET STRING }

tt TT ::= {a 77,b {"kalle","kula"}}
```

The value defined here could be used in several ways. Firstly, it could be used as the value in some DEFAULT component:

```
SS ::= SET {
    s OBJECT IDENTIFIER,
    val TT DEFAULT tt }
```

It could also be used from inside an Erlang program. If the above ASN.1 code was defined in ASN.1 module `Values`, then the ASN.1 value `tt` can be reached from Erlang as a function call to `'Values':tt()` as in the example below.

```
1> Val = 'Values':tt().
{'TT',77,["kalle","kula"]}
2> {ok,Bytes} = 'Values':encode('TT',Val).
{ok,<<48,18,128,1,77,161,13,4,5,107,97,108,108,101,4,4,
    107,117,108,97>>}
4> 'Values':decode('TT',Bytes).
{ok,{'TT',77,["kalle","kula"]}}
5>
```

The above example shows that a function is generated by the compiler that returns a valid Erlang representation of the value, even though the value is of a complex type.

Furthermore, there is a macro generated for each value in the .hrl file. So, the defined value `tt` can also be extracted by `?tt` in application code.

1.1.9 Macros

MACRO is not supported as the type is no longer part of the ASN.1 standard.

1.1.10 ASN.1 Information Objects (X.681)

Information Object Classes, Information Objects and Information Object Sets (in the following called classes, objects and object sets respectively) are defined in the standard definition []. In the following only a brief explanation is given.

These constructs makes it possible to define open types, i.e. values of that type can be of any ASN.1 type. It is also possible to define relationships between different types and values, since classes can hold types, values, objects, object sets and other classes in its fields. An Information Object Class may be defined in ASN.1 as:

```
GENERAL-PROCEDURE ::= CLASS {
```

```

    &Message,
    &Reply          OPTIONAL,
    &Error          OPTIONAL,
    &id             PrintableString UNIQUE
}
WITH SYNTAX {
    NEW MESSAGE      &Message
    [REPLY           &Reply]
    [ERROR           &Error]
    ADDRESS          &id
}

```

An object is an instance of a class and an object set is a set containing objects of one specified class. A definition may look like below.

The object `object1` is an instance of the CLASS GENERAL-PROCEDURE and has one type field and one fixed type value field. The object `object2` also has an OPTIONAL field ERROR, which is a type field.

```

object1 GENERAL-PROCEDURE ::= {
    NEW MESSAGE      PrintableString
    ADDRESS          "home"
}

object2 GENERAL-PROCEDURE ::= {
    NEW MESSAGE INTEGER
    ERROR INTEGER
    ADDRESS "remote"
}

```

The field ADDRESS is a UNIQUE field. Objects in an object set must have unique values in their UNIQUE field, as in GENERAL-PROCEDURES:

```

GENERAL-PROCEDURES GENERAL-PROCEDURE ::= {
    object1 | object2
}

```

One can not encode a class, object or object set, only referring to it when defining other ASN.1 entities. Typically one refers to a class and to object sets by table constraints and component relation constraints [] in ASN.1 types, as in:

```

StartMessage ::= SEQUENCE {
    msgId GENERAL-PROCEDURE.&id ({GENERAL-PROCEDURES}),
    content GENERAL-PROCEDURE.&Message ({GENERAL-PROCEDURES}{@msgId}),
}

```

In the type `StartMessage` the constraint following the `content` field tells that in a value of type `StartMessage` the value in the `content` field must come from the same object that is chosen by the `msgId` field.

So, the value `#'StartMessage' {msgId="home", content="Any Printable String"}` is legal to encode as a `StartMessage` value, while the value `#'StartMessage' {msgId="remote", content="Some String"}` is illegal since the constraint in `StartMessage` tells that when you have chosen a value from a specific object in the object set `GENERAL-PROCEDURES` in the `msgId` field you have to choose a value from that same object in the `content` field too. In this second case it should have been any `INTEGER` value.

1.2 Specialized Decodes

StartMessage can in the content field be encoded with a value of any type that an object in the GENERAL-PROCEDURES object set has in its NEW MESSAGE field. This field refers to a type field &Message in the class. The msgId field is always encoded as a PrintableString, since the field refers to a fixed type in the class.

In practice, object sets are usually declared to be extensible so so that more objects can be added to the set later. Extensibility is indicated like this:

```
GENERAL-PROCEDURES GENERAL-PROCEDURE ::= {  
    object1 | object2, ...}
```

When decoding a type that uses an extensible set constraint, there is always the possibility that the value in the UNIQUE field is unknown (i.e. the type has been encoded with a later version of the ASN.1 specification). When that happens, the unencoded data will be returned wrapped in a tuple like this:

```
{asn1_OPENTYPE, Binary}
```

where Binary is an Erlang binary that contains the encoded data. (If the option legacy_erlang_types has been given, just the binary will be returned.)

1.1.11 Parameterization (X.683)

Parameterization, which is defined in the standard [], can be used when defining types, values, value sets, information object classes, information objects or information object sets. A part of a definition can be supplied as a parameter. For instance, if a Type is used in a definition with certain purpose, one want the type-name to express the intention. This can be done with parameterization.

When many types (or another ASN.1 entity) only differs in some minor cases, but the structure of the types are similar, only one general type can be defined and the differences may be supplied through parameters.

One example of use of parameterization is:

```
General{Type} ::= SEQUENCE  
{  
    number    INTEGER,  
    string    Type  
}  
  
T1 ::= General{PrintableString}  
T2 ::= General{BIT STRING}
```

An example of a value that can be encoded as type T1 is {12,"hello"}.

Note that the compiler does not generate encode/decode functions for parameterized types, but only for the instances of the parameterized types. Therefore, if a file contains the types General{ }, T1 and T2 above, encode/decode functions will only be generated for T1 and T2.

1.2 Specialized Decodes

When performance is of highest priority and one is interested in a limited part of the ASN.1 encoded message, before one decide what to do with the rest of it, one may want to decode only this small part. The situation may be a server that has to decide to which addressee it will send a message. The addressee may be interested in the entire message, but

the server may be a bottleneck that one want to spare any unnecessary load. Instead of making two *complete decodes* (the normal case of decode), one in the server and one in the addressee, it is only necessary to make one *specialized decode* (in the server) and another complete decode (in the addressee). The following specialized decodes *exclusive decode* and *selected decode* support to solve this and similar problems.

So far this functionality is only provided when using the optimized BER_BIN version, that is when compiling with the options `ber_bin` and `optimize`. It does also work using the `nif` option. We have no intent to make this available on the default BER version, but maybe in the PER_BIN version (`per_bin`).

1.2.1 Exclusive Decode

The basic idea with exclusive decode is that you specify which parts of the message you want to exclude from being decoded. These parts remain encoded and are returned in the value structure as binaries. They may be decoded in turn by passing them to a certain `decode_part/2` function. The performance gain is high when the message is large and you can do an exclusive decode and later on one or several decodes of the parts or a second complete decode instead of two or more complete decodes.

How To Make It Work

In order to make exclusive decode work you have to do the following:

- First, decide the name of the function for the exclusive decode.
- Second, write instructions that must consist of the name of the exclusive decode function, the name of the ASN.1 specification and a notation that tells which parts of the message structure will be excluded from decode. These instructions shall be included in a configuration file.
- Third, compile with the additional option `asn1config`. The compiler searches for a configuration file with the same name as the ASN.1 spec but with the extension `.asn1config`. This configuration file is not the same as used for compilation of a set of files. See section *Writing an Exclusive Decode Instruction*.

User Interface

The run-time user interface for exclusive decode consists of two different functions. First, the function for an exclusive decode, whose name the user decides in the configuration file. Second, the compiler generates a `decode_part/2` function when exclusive decode is chosen. This function decodes the parts that were left undecoded during the exclusive decode. Both functions are described below.

If the exclusive decode function has for example got the name `decode_exclusive` and an ASN.1 encoded message `Bin` shall be exclusive decoded, the call is:

```
{ok,Excl_Message} = 'MyModule':decode_exclusive(Bin)
```

The result `Excl_Message` has the same structure as an complete decode would have, except for the parts of the top-type that were not decoded. The undecoded parts will be on their place in the structure on the format `{Type_Key, Undecoded_Value}`.

Each undecoded part that shall be decoded must be fed into the `decode_part/2` function, like:

```
{ok,Part_Message} = 'MyModule':decode_part(Type_Key,Undecoded_Value)
```

Writing an Exclusive Decode Instruction

This instruction is written in the configuration file on the format:

1.2 Specialized Decodes

```
Exclusive_Decode_Instruction = {exclusive_decode,{Module_Name,Decode_Instructions}}.  
  
Module_Name = atom()  
  
Decode_Instructions = [Decode_Instruction]+  
  
Decode_Instruction = {Exclusive_Decode_Function_Name,Type_List}  
  
Exclusive_Decode_Function_Name = atom()  
  
Type_List = [Top_Type,Element_List]  
  
Element_List = [Element]+  
  
Element = {Name,parts} |  
          {Name,undecoded} |  
          {Name,Element_List}  
  
Top_Type = atom()  
  
Name = atom()
```

Observe that the instruction must be a valid Erlang term ended by a dot.

In the `Type_List` the "path" from the top type to each undecoded sub-components is described. The top type of the path is an atom, the name of it. The action on each component/type that follows will be described by one of `{Name,parts}`, `{Name,undecoded}`, `{Name,Element_List}`

The use and effect of the actions are:

- `{Name,undecoded}` Tells that the element will be left undecoded during the exclusive decode. The type of Name may be any ASN.1 type. The value of element Name will be returned as a tuple, as mentioned *above*, in the value structure of the top type.
- `{Name,parts}` The type of Name may be one of SEQUENCE OF or SET OF. The action implies that the different components of Name will be left undecoded. The value of Name will be returned as a tuple, as *above*, where the second element is a list of binaries. That is because the representation of a SEQUENCE OF/ SET OF in Erlang is a list of its internal type. Any of the elements of this list or the entire list can be decoded by the `decode_part` function.
- `{Name,Element_List}` This action is used when one or more of the sub-types of Name will be exclusive decoded.

Name in the actions above may be a component name of a SEQUENCE or a SET or a name of an alternative in a CHOICE.

Example

In the examples below we use the definitions from the following ASN.1 spec:

```
GUI DEFINITIONS AUTOMATIC TAGS ::=
BEGIN

Action ::= SEQUENCE
{
    number    INTEGER DEFAULT 15,
    handle    [0] Handle DEFAULT {number 12, on TRUE}
}
```

```

Key ::= [11] EXPLICIT Button
Handle ::= [12] Key
Button ::= SEQUENCE
{
    number INTEGER,
    on BOOLEAN
}

Window ::= CHOICE
{
    vsn INTEGER,
    status E
}

Status ::= SEQUENCE
{
    state INTEGER,
    buttonList SEQUENCE OF Button,
    enabled BOOLEAN OPTIONAL,
    actions CHOICE {
        possibleActions SEQUENCE OF Action,
        noOfActions INTEGER
    }
}

END

```

If Button is a top type and we want to exclude component number from decode the Type_List in the instruction in the configuration file will be ['Button', [{number, undecoded}]]. If we call the decode function decode_Button_exclusive the Decode_Instruction will be {decode_Button_exclusive, ['Button', [{number, undecoded}]]}.

We also have another top type Window whose sub component actions in type Status and the parts of component buttonList shall be left undecoded. For this type we name the function decode__Window_exclusive. The whole Exclusive_Decompile_Instruction configuration is as follows:

```

{exclusive_decode, {'GUI',
  [{decode_Window_exclusive, ['Window', [{status, [{buttonList, parts}, {actions, undecoded}]}]}],
  {decode_Button_exclusive, ['Button', [{number, undecoded}]}]}]}.

```

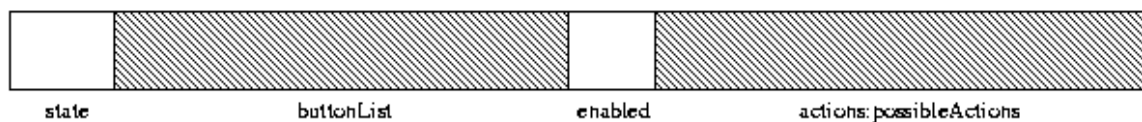


Figure 2.1: Figure symbolizes the bytes of a Window:status message. The components buttonList and actions are excluded from decode. Only state and enabled are decoded when decode__Window_exclusive is called.

Compiling GUI.asn including the configuration file is done like:

```

unix> erlc -bber_bin +optimize +asn1config GUI.asn

```

1.2 Specialized Decodes

```
erlang> asn1ct:compile('GUI',[ber_bin,optimize,asn1config]).
```

The module can be used like:

```
1> Button_Msg = {'Button',123,true}.
{'Button',123,true}
2> {ok,Button_Bytes} = 'GUI':encode('Button',Button_Msg).
{ok,[<<48>>,
    [6],
    [<<128>>,
    [1],
    123],
    [<<129>>,
    [1],
    255]]}
3> {ok,Exclusive_Msg_Button} = 'GUI':decode_button_exclusive(list_to_binary(Button_Bytes)).
{ok,{'Button',{'Button_number',<<28,1,123>>},
    true}}
4> 'GUI':decode_part('Button_number',<<128,1,123>>).
{ok,123}
5> Window_Msg =
{'Window',{status,{'Status',35,
    [{'Button',3,true},
    {'Button',4,false},
    {'Button',5,true},
    {'Button',6,true},
    {'Button',7,false},
    {'Button',8,true},
    {'Button',9,true},
    {'Button',10,false},
    {'Button',11,true},
    {'Button',12,true},
    {'Button',13,false},
    {'Button',14,true}],
    false,
    {possibleActions,[{'Action',16,{'Button',17,true}}]}]}},
{'Window',{status,{'Status',35,
    [{'Button',3,true},
    {'Button',4,false},
    {'Button',5,true},
    {'Button',6,true},
    {'Button',7,false},
    {'Button',8,true},
    {'Button',9,true},
    {'Button',10,false},
    {'Button',11,true},
    {'Button',12,true},
    {'Button',13,false},
    {'Button',14,true}],
    false,
    {possibleActions,[{'Action',16,{'Button',17,true}}]}}}}
6> {ok,Window_Bytes}='GUI':encode('Window',Window_Msg).
{ok,[<<161>>,
    [127],
    [<<128>>, ...
7>
8> {ok,{status,{'Status',Int,{Type_Key_SeqOf,Val_SEQOF},
    BoolOpt,{Type_Key_Choice,Val_Choice}}}}=
'GUI':decode_window_status_exclusive(list_to_binary(Window_Bytes)).
{ok,{status,{'Status',35,
```

```

        {'Status_buttonList', [<<48,6,128,1,3,129,1,255>>,
                                <<48,6,128,1,4,129,1,0>>,
                                <<48,6,128,1,5,129,1,255>>,
                                <<48,6,128,1,6,129,1,255>>,
                                <<48,6,128,1,7,129,1,0>>,
                                <<48,6,128,1,8,129,1,255>>,
                                <<48,6,128,1,9,129,1,255>>,
                                <<48,6,128,1,10,129,1,0>>,
                                <<48,6,128,1,11,129,1,255>>,
                                <<48,6,128,1,12,129,1,255>>,
                                <<48,6,128,1,13,129,1,0>>,
                                <<48,6,128,1,14,129,1,255>>]},
        false,
        {'Status_actions',
         <<163,21,160,19,48,17,2,1,16,160,12,172,10,171,8,48,6,128,1,...>>}}}}
10> 'GUI':decode_part(Type_Key_SeqOf,Val_SEQOF).
{ok,[{'Button',3,true},
      {'Button',4,false},
      {'Button',5,true},
      {'Button',6,true},
      {'Button',7,false},
      {'Button',8,true},
      {'Button',9,true},
      {'Button',10,false},
      {'Button',11,true},
      {'Button',12,true},
      {'Button',13,false},
      {'Button',14,true}]}
11> 'GUI':decode_part(Type_Key_SeqOf,hd(Val_SEQOF)).
{ok,{'Button',3,true}}
12> 'GUI':decode_part(Type_Key_Choice,Val_Choice).
{ok,{possibleActions,[{'Action',16,{'Button',17,true}}]}}

```

1.2.2 Selective Decode

This specialized decode decodes one single subtype of a constructed value. It is the fastest method to extract one sub value. The typical use of this decode is when one want to inspect, for instance a version number, to be able to decide what to do with the entire value. The result is returned as {ok, Value} or {error, Reason}.

How To Make It Work

The following steps are necessary:

- Write instructions in the configuration file. Including the name of a user function, the name of the ASN.1 specification and a notation that tells which part of the type will be decoded.
- Compile with the additional option `asn1config`. The compiler searches for a configuration file with the same name as the ASN.1 spec but with the extension `.asn1config`. In the same file you can provide configuration specs for exclusive decode as well. The generated Erlang module has the usual functionality for encode/decode preserved and the specialized decode functionality added.

User Interface

The only new user interface function is the one provided by the user in the configuration file. You can invoke that function by the `ModuleName:FunctionName` notation.

So, if you have the following spec `{selective_decode, {'ModuleName', [{selected_decode_Window, TypeList}]}}` in the con-file, you do the selective decode by `{ok, Result} = 'ModuleName':selected_decode_Window(EncodedBinary)`.

Writing a Selective Decode Instruction

It is possible to describe one or many selective decode functions in a configuration file, you have to use the following notation:

```
Selective_Decode_Instruction = {selective_decode,{Module_Name,Decode_Instructions}}.  
  
Module_Name = atom()  
  
Decode_Instructions = [Decode_Instruction]+  
  
Decode_Instruction = {Selective_Decode_Function_Name,Type_List}  
  
Selective_Decode_Function_Name = atom()  
  
Type_List = [Top_Type|Element_List]  
  
Element_List = Name|List_Selector  
  
Name = atom()  
  
List_Selector = [integer()]
```

Observe that the instruction must be a valid Erlang term ended by a dot.

The `Module_Name` is the same as the name of the ASN.1 spec, but without the extension. A `Decode_Instruction` is a tuple with your chosen function name and the components from the top type that leads to the single type you want to decode. Notice that you have to choose a name of your function that will not be the same as any of the generated functions. The first element of the `Type_List` is the top type of the encoded message. In the `Element_List` it is followed by each of the component names that leads to selected type. Each of the names in the `Element_List` must be constructed types except the last name, which can be any type.

The `List_Selector` makes it possible to choose one of the encoded components in a SEQUENCE OF/ SET OF. It is also possible to go further in that component and pick a sub type of that to decode. So in the `Type_List`: `['Window',status,buttonList,[1],number]` the component `buttonList` has to be a SEQUENCE OF or SET OF type. In this example component `number` of the first of the encoded elements in the SEQUENCE OF `buttonList` is selected. This apply on the ASN.1 spec *above*.

Another Example

In this example we use the same ASN.1 spec as *above*. A valid selective decode instruction is:

```
{selective_decode,  
  {'GUI',  
    [{selected_decode_Window1,  
      ['Window',status,buttonList,  
        [1],  
        number]],  
    {selected_decode_Action,  
      ['Action',handle,number]],  
    {selected_decode_Window2,  
      ['Window',  
        status,  
        actions,  
        possibleActions,  
        [1],  
        handle,number]]}}}.  
}
```

The first `Decode_Instruction`, `{selected_decode_Window1,['Window',status,buttonList,[1],number]}` is commented in the previous section. The instruction `{selected_decode_Action,['Action',handle,number]}` picks the component number in the `handle` component of the type `Action`. If we have the value `ValAction = {'Action',17,{'Button',4711,false}}` the internal value 4711 should be picked by `selected_decode_Action`. In an Erlang terminal it looks like:

```
ValAction = {'Action',17,{'Button',4711,false}}.
{'Action',17,{'Button',4711,false}}
7> {ok,Bytes}= 'GUI':encode('Action',ValAction).
...
8> BinBytes = list_to_binary(Bytes).
<<48,18,2,1,17,160,13,172,11,171,9,48,7,128,2,18,103,129,1,0>>
9> 'GUI':selected_decode_Action(BinBytes).
{ok,4711}
10>
```

The third instruction, `['Window',status,actions,possibleActions,[1],handle,number]`, which is a little more complicated,

- starts with type *Window*.
- Picks component *status* of *Window* that is of type *Status*.
- Then takes component *actions* of type *Status*.
- Then *possibleActions* of the internal defined CHOICE type.
- Thereafter it goes into the first component of the SEQUENCE OF by *[1]*. That component is of type *Action*.
- The instruction next picks component *handle*.
- And finally component *number* of the type *Button*.

The following figures shows which components are in the `TypeList` `['Window',status,actions,possibleActions,[1],handle,number]`. And which part of a message that will be decoded by `selected_decode_Window2`.

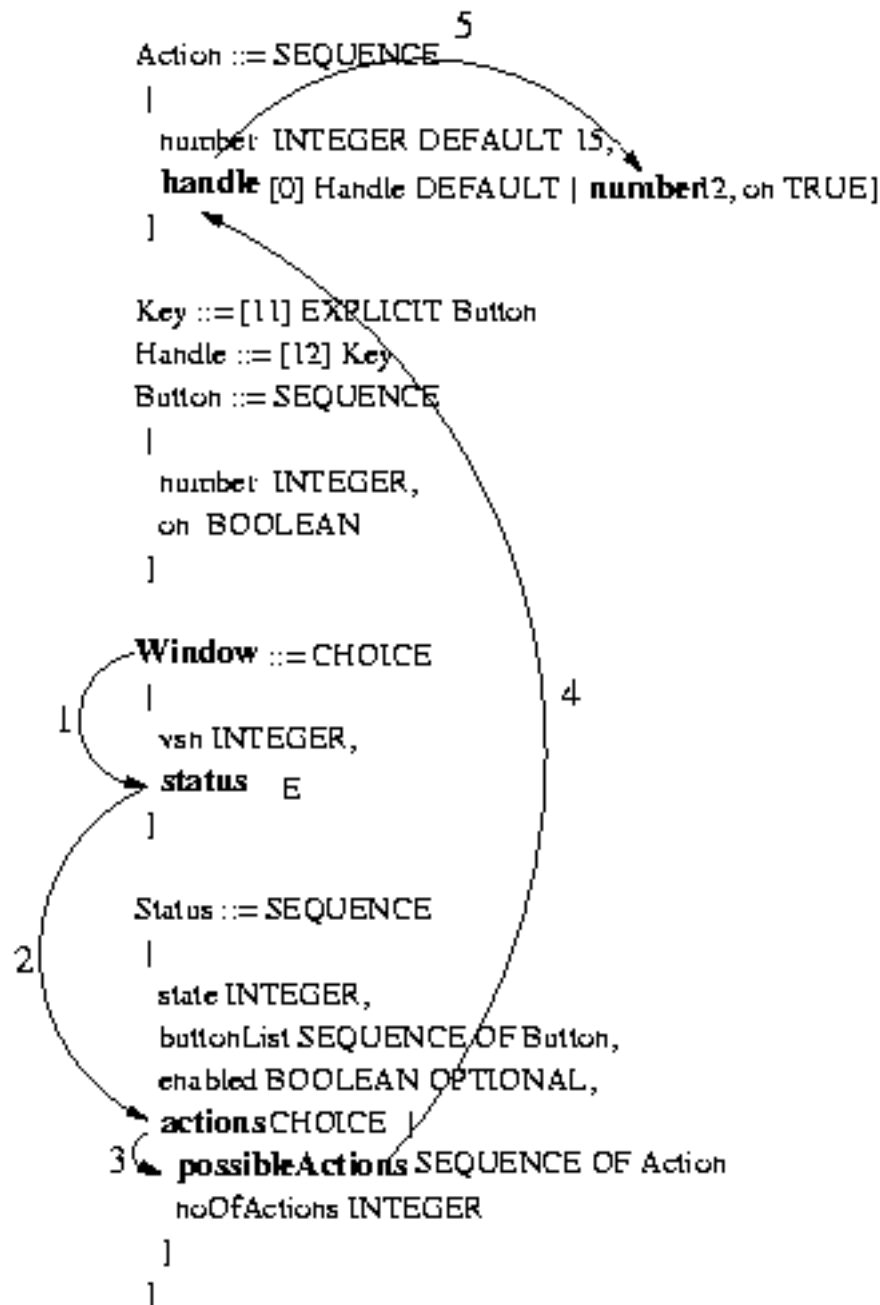


Figure 2.2: The elements specified in the config file for selective decode of a sub-value in a Window message



Figure 2.3: Figure symbolizes the bytes of a Window:status message. Only the marked element is decoded when `selected_decode_Window2` is called.

With the following example you can examine that both `selected_decode_Window2` and `selected_decode_Window1` decodes the intended sub-value of the value `Val`

```
1> Val = { 'Window', {status, { 'Status', 12,
    [ { 'Button', 13, true },
      { 'Button', 14, false },
      { 'Button', 15, true },
      { 'Button', 16, false } ],
    true,
    {possibleActions, [ { 'Action', 17, { 'Button', 18, false } },
                        { 'Action', 19, { 'Button', 20, true } },
                        { 'Action', 21, { 'Button', 22, false } } ] } } } }
2> {ok, Bytes} = 'GUI':encode( 'Window', Val ).
...
3> Bin = list_to_binary(Bytes).
<<161,101,128,1,12,161,32,48,6,128,1,13,129,1,255,48,6,128,1,14,129,1,0,48,6,128,1,15,129,...>>
4> 'GUI':selected_decode_Window1(Bin).
{ok,13}
5> 'GUI':selected_decode_Window2(Bin).
{ok,18}
```

Observe that the value feed into the selective decode functions must be a binary.

1.2.3 Performance

To give an indication on the possible performance gain using the specialized decodes, some measures have been performed. The relative figures in the outcome between selective, exclusive and complete decode (the normal case) depends on the structure of the type, the size of the message and on what level the selective and exclusive decodes are specified.

1.2 Specialized Decodes

ASN.1 Specifications, Messages and Configuration

The specs *GUI* and **MEDIA-GATEWAY-CONTROL** was used in the test.

For the GUI spec the configuration looked like:

```
{selective_decode,
  {'GUI',
    [{selected_decode_Window1,
      ['Window',
        status,buttonList,
        [1],
        number]}],
    {selected_decode_Window2,
      ['Window',
        status,
        actions,
        possibleActions,
        [1],
        handle,number]]}].
{exclusive_decode,
  {'GUI',
    [{decode_Window_status_exclusive,
      ['Window',
        [{status,
          [{buttonList,parts},
            {actions,undecoded}]]]]}].
```

The MEDIA-GATEWAY-CONTROL configuration was:

```
{exclusive_decode,
  {'MEDIA-GATEWAY-CONTROL',
    [{decode_MegacoMessage_exclusive,
      ['MegacoMessage',
        [{authHeader,undecoded},
          {mess,
            [{mId,undecoded},
              {messageBody,undecoded}]]]]]}].
{selective_decode,
  {'MEDIA-GATEWAY-CONTROL',
    [{decode_MegacoMessage_selective,
      ['MegacoMessage',mess,version]]}].
```

The corresponding values were:

```
{'Window', {status, {'Status',12,
  [{ 'Button',13,true},
    { 'Button',14,false},
    { 'Button',15,true},
    { 'Button',16,false},
    { 'Button',13,true},
    { 'Button',14,false},
    { 'Button',15,true},
    { 'Button',16,false},
    { 'Button',13,true},
    { 'Button',14,false},
```

```

        {'Button',15,true},
        {'Button',16,false}]],
    true,
    {possibleActions,
      [{ 'Action',17,{'Button',18,false}},
        { 'Action',19,{'Button',20,true}},
        { 'Action',21,{'Button',22,false}},
        { 'Action',17,{'Button',18,false}},
        { 'Action',19,{'Button',20,true}},
        { 'Action',21,{'Button',22,false}},
        { 'Action',17,{'Button',18,false}},
        { 'Action',19,{'Button',20,true}},
        { 'Action',21,{'Button',22,false}},
        { 'Action',17,{'Button',18,false}},
        { 'Action',19,{'Button',20,true}},
        { 'Action',21,{'Button',22,false}},
        { 'Action',17,{'Button',18,false}},
        { 'Action',19,{'Button',20,true}},
        { 'Action',21,{'Button',22,false}},
        { 'Action',17,{'Button',18,false}},
        { 'Action',19,{'Button',20,true}},
        { 'Action',21,{'Button',22,false}}]}]}]}

{ 'MegacoMessage',asn1_NOVALUE,
  { 'Message',1,
    {ip4Address,
      { 'IP4Address',[125,125,125,111],55555}},
    {transactions,
      [{transactionReply,
        { 'TransactionReply',50007,asn1_NOVALUE,
          {actionReplies,
            [{ 'ActionReply',0,asn1_NOVALUE,asn1_NOVALUE,
              [{auditValueReply,{auditResult,{ 'AuditResult',
                { 'TerminationID',[],[255,255,255]}},
                [{mediaDescriptor,
                  { 'MediaDescriptor',asn1_NOVALUE,
                    {multiStream,
                      [{ 'StreamDescriptor',1,
                        { 'StreamParms',
                          { 'LocalControlDescriptor',
                            sendRecv,
                            asn1_NOVALUE,
                            asn1_NOVALUE,
                            [{ 'PropertyParm',
                              [0,11,0,7],
                              [[52,48]],
                              asn1_NOVALUE]}]},
                        { 'LocalRemoteDescriptor',
                          [{ { 'PropertyParm',
                              [0,0,176,1],
                              [[48]],
                              asn1_NOVALUE},
                            { 'PropertyParm',
                              [0,0,176,8],
                              [[73,78,32,73,80,52,32,49,50,53,46,49,
                                50,53,46,49,50,53,46,49,49,49]],
                              asn1_NOVALUE},
                            { 'PropertyParm',
                              [0,0,176,15],
                              [[97,117,100,105,111,32,49,49,49,49,32,
                                82,84,80,47,65,86,80,32,32,52]],
                              asn1_NOVALUE},
                            { 'PropertyParm',
                              [0,0,176,12],

```

The size of the encoded values was 458 bytes for GUI and 464 bytes for MEDIA-GATEWAY-CONTROL.

34 | Ericsson AB. All Rights Reserved.: ASN.1

selected_decode_Window2/1	8910666	selective	GUI	15.1
decode_Window_status-exclusive/1	1251878	exclusive	GUI	21.3
decode/2	5889197	complete	GUI	100

Table 2.1: Results of complete, exclusive and selective decode

Another interesting question is what the relation is between a complete decode, an exclusive decode followed by decode_part of the excluded parts and a selective decode followed by a complete decode. Some situations may be compared to this simulation, e.g. inspect a sub-value and later on look at the entire value. The following table shows figures from this test. The number of loops and time unit is the same as in the previous test.

<i>Actions</i>	<i>Function</i>	<i>Time(microseconds)</i>	<i>ASN.1 spec</i>	<i>% of time vs. complete decode</i>
complete	decode/2	4507457	MEDIA-GATEWAY-CONTROL	100
selective and complete	decode_MegacoMessage_selective/1	4881502	MEDIA-GATEWAY-CONTROL	108.3
exclusive and decode_part	decode_MegacoMessage-exclusive/1	5481034	MEDIA-GATEWAY-CONTROL	112.3
complete	decode/2	5889197	GUI	100
selective and complete	selected_decode_Window1/1	6337636	GUI	107.6
selective and complete	selected_decode_Window2/1	6795319	GUI	115.4
exclusive and decode_part	decode_Window_status-exclusive/1	6249200	GUI	106.1

Table 2.2: Results of complete, exclusive + decode_part and selective + complete decodes

Other ASN.1 types and values can differ much from these figures. Therefore it is important that you, in every case where you intend to use either of these decodes, perform some tests that shows if you will benefit your purpose.

Comments

Generally speaking the gain of selective and exclusive decode in advance of complete decode is greater the bigger value and the less deep in the structure you have to decode. One should also prefer selective decode instead of exclusive decode if you are interested in just one single sub-value.

Another observation is that the exclusive decode followed by `decode_part` decodes is very attractive if the parts will be sent to different servers for decoding or if one in some cases not is interested in all parts.

The fastest selective decode are when the decoded type is a primitive type and not so deep in the structure of the top type. The `selected_decode_window2` decodes a big constructed value, which explains why this operation is relatively slow.

It may vary from case to case which combination of selective/complete decode or exclusive/part decode is the fastest.

2 Reference Manual

The *Asn1* application contains modules with compile-time and run-time support for ASN.1.

asn1ct

Erlang module

The ASN.1 compiler takes an ASN.1 module as input and generates a corresponding Erlang module which can encode and decode the data-types specified. Alternatively the compiler takes a specification module (see below) specifying all input modules and generates one module with encode/decode functions. There are also some generic functions which can be used in during development of applications which handles ASN.1 data (encoded as BER or PER).

Note:

By default in OTP 17, the representation of the BIT STRING and OCTET STRING types as Erlang terms have changed. BIT STRING values are now Erlang bitstrings and OCTET STRING values are binaries. Also, an undecoded open type will now be wrapped in a `asn1_OPEN_TYPE` tuple. For details see *BIT STRING*, *OCTET STRING*, and *ASN.1 Information Objects* in User's Guide.

To revert to the old representation of the types, use the `legacy_erlang_types` option.

Note:

In R16, the options have been simplified. The back-end is chosen using one of the options `ber`, `per`, or `uper`. The options `optimize`, `nif`, and `driver` options are no longer necessary (and the ASN.1 compiler will print a warning if they are used). The options `ber_bin`, `per_bin`, and `uper_bin` options will still work, but will print a warning.

Another change in R16 is that the generated `encode/2` function always returns a binary. The `encode/2` function for the BER back-end used to return an iolist.

Exports

```
compile(Asn1module) -> ok | {error, Reason}
compile(Asn1module, Options) -> ok | {error, Reason}
```

Types:

```
Asn1module = atom() | string()
Options = [Option | OldOption]
Option = ber | per | uper | der | compact_bit_string | legacy_bit_string
| legacy_erlang_types | noobj | {n2n, EnumTypeName} | {outdir,
Dir} | {i, IncludeDir} | asn1config | undec_rest | no_ok_wrapper |
{macro_name_prefix, Prefix} | {record_name_prefix, Prefix} | verbose |
warnings_as_errors
OldOption = ber | per
Reason = term()
Prefix = string()
```

Compiles the ASN.1 module `Asn1module` and generates an Erlang module `Asn1module.erl` with encode and decode functions for the types defined in `Asn1module`. For each ASN.1 value defined in the module an Erlang function which returns the value in Erlang representation is generated.

If `Asn1module` is a filename without extension first ".asn1" is assumed, then ".asn" and finally ".py" (to be compatible with the old ASN.1 compiler). Of course `Asn1module` can be a full pathname (relative or absolute) including filename with (or without) extension.

If one wishes to compile a set of `Asn1` modules into one Erlang file with encode/decode functions one has to list all involved files in a configuration file. This configuration file must have a double extension ".set.asn", (".asn" can alternatively be ".asn1" or ".py"). The input files' names must be listed, within quotation marks (""), one at each row in the file. If the input files are `File1.asn`, `File2.asn` and `File3.asn` the configuration file shall look like:

```
File1.asn
File2.asn
File3.asn
```

The output files will in this case get their names from the configuration file. If the configuration file has the name `SetOfFiles.set.asn` the name of the output files will be `SetOfFiles.hrl`, `SetOfFiles.erl` and `SetOfFiles.asnldb`.

Sometimes in a system of ASN.1 modules there are different default tag modes, e.g. AUTOMATIC, IMPLICIT or EXPLICIT. The multi file compilation resolves the default tagging as if the modules were compiled separately.

Another unwanted effect that may occur in multi file compilation is name collisions. The compiler solves this problem in two ways: If the definitions are identical then the output module keeps only one definition with the original name. But if definitions only have same name and differs in the definition, then they will be renamed. The new names will be the definition name and the original module name concatenated.

If any name collision have occurred the compiler reports a "NOTICE: ..." message that tells if a definition was renamed, and the new name that must be used to encode/decode data.

`Options` is a list with options specific for the `asn1` compiler and options that are applied to the Erlang compiler. The latter are those that not is recognized as `asn1` specific. Available options are:

`ber` | `per` | `uper`

The encoding rule to be used. The supported encoding rules are BER (Basic Encoding Rules), PER aligned (Packed Encoding Rules) and PER unaligned. If the encoding rule option is omitted `ber` is the default.

The generated Erlang module always gets the same name as the ASN.1 module and as a consequence of this only one encoding rule per ASN.1 module can be used at runtime.

`der`

By this option the Distinguished Encoding Rules (DER) is chosen. DER is regarded as a specialized variant of the BER encoding rule, therefore the `der` option only makes sense together with the `ber` option. This option sometimes adds sorting and value checks when encoding, which implies a slower encoding. The decoding routines are the same as for `ber`.

`compact_bit_string`

The BIT STRING type will be decoded to the "compact notation". *This option is not recommended for new code.*

For details see *BIT STRING type section in the Users Guide*.

This option implies the `legacy_erlang_types` option.

`legacy_bit_string`

The BIT STRING type will be decoded to the legacy format, i.e. a list of zeroes and ones. *This option is not recommended for new code.*

For details see *BIT STRING type section in the Users Guide*.

This option implies the `legacy_erlang_types` option.

legacy_erlang_types

Use the same Erlang types to represent BIT STRING and OCTET STRING as in R16. For details see *BIT STRING* and *OCTET STRING* in User's Guide.

This option is not recommended for new code.

{n2n, EnumTypeName}

Tells the compiler to generate functions for conversion between names (as atoms) and numbers and vice versa for the EnumTypeName specified. There can be multiple occurrences of this option in order to specify several type names. The type names must be declared as ENUMERATIONS in the ASN.1 spec. If the EnumTypeName does not exist in the ASN.1 spec the compilation will stop with an error code. The generated conversion functions are named name2num_EnumTypeName/1 and num2name_EnumTypeName/1.

noobj

Do not compile (i.e do not produce object code) the generated .erl file. If this option is omitted the generated Erlang module will be compiled.

{i, IncludeDir}

Adds IncludeDir to the search-path for .asn1db and asn1 source files. The compiler tries to open a .asn1db file when a module imports definitions from another ASN.1 module. If no .asn1db file is found the asn1 source file is parsed. Several {i, IncludeDir} can be given.

{outdir, Dir}

Specifies the directory Dir where all generated files shall be placed. If omitted the files are placed in the current directory.

asn1config

When one of the specialized decodes, exclusive or selective decode, is wanted one has to give instructions in a configuration file. The option asn1config enables specialized decodes and takes the configuration file, which has the same name as the ASN.1 spec but with extension .asn1config, in concern.

The instructions for exclusive decode must follow the *instruction and grammar in the User's Guide*.

You can also find the instructions for selective decode in the *User's Guide*.

undec_rest

A buffer that holds a message, being decoded may also have some following bytes. Now it is possible to get those following bytes returned together with the decoded value. If an asn1 spec is compiled with this option a tuple {ok, Value, Rest} is returned. Rest may be a list or a binary. Earlier versions of the compiler ignored those following bytes.

no_ok_wrapper

If this option is given, the generated encode/2 and decode/2 functions will not wrap a successful return value in an {ok, ...} tuple. If any error occurs, there will be an exception.

{macro_name_prefix, Prefix}

All macro names generated by the compiler are prefixed with Prefix. This is useful when multiple protocols that contains macros with identical names are included in a single module.

{record_name_prefix, Prefix}

All record names generated by the compiler are prefixed with Prefix. This is useful when multiple protocols that contains records with identical names are included in a single module.

verbose

Causes more verbose information from the compiler describing what it is doing.

warnings_as_errors

Causes warnings to be treated as errors.

Any additional option that is applied will be passed to the final step when the generated .erl file is compiled.

The compiler generates the following files:

- Asn1module.hrl (if any SET or SEQUENCE is defined)
- Asn1module.erl the Erlang module with encode, decode and value functions.
- Asn1module.asn1db intermediate format used by the compiler when modules IMPORTS definitions from each other.

encode(Module, Type, Value) -> {ok, Bytes} | {error, Reason}

Types:

```
Module = Type = atom()  
Value = term()  
Bytes = binary()  
Reason = term()
```

Encodes Value of Type defined in the ASN.1 module Module. To get as fast execution as possible the encode function only performs rudimentary tests that the input Value is a correct instance of Type. The length of strings is for example not always checked. Returns {ok, Bytes} if successful or {error, Reason} if an error occurred.

This function is deprecated. Use Module:encode(Type, Value) instead.

decode(Module, Type, Bytes) -> {ok, Value} | {error, Reason}

Types:

```
Module = Type = atom()  
Value = Reason = term()  
Bytes = binary()
```

Decodes Type from Module from the binary Bytes. Returns {ok, Value} if successful.

This function is deprecated. Use Module:decode(Type, Bytes) instead.

value(Module, Type) -> {ok, Value} | {error, Reason}

Types:

```
Module = Type = atom()  
Value = term()  
Reason = term()
```

Returns an Erlang term which is an example of a valid Erlang representation of a value of the ASN.1 type Type. The value is a random value and subsequent calls to this function will for most types return different values.

test(Module) -> ok | {error, Reason}

test(Module, Type | Options) -> ok | {error, Reason}

test(Module, Type, Value | Options) -> ok | {error, Reason}

Types:

```
Module = Type = atom()  
Value = term()  
Options = [{i, IncludeDir}]
```

Reason = term()

Performs a test of encode and decode of types in `Module`. The generated functions are called by this function. This function is useful during test to secure that the generated encode and decode functions and the general runtime support work as expected.

- `test/1` iterates over all types in `Module`.
- `test/2` tests type `Type` with a random value.
- `test/3` tests type `Type` with `Value`.

Schematically the following happens for each type in the module:

```
{ok, Value} = asn1ct:value(Module, Type),  
{ok, Bytes} = asn1ct:encode(Module, Type, Value),  
{ok, Value} = asn1ct:decode(Module, Type, Bytes).
```

The `test` functions utilizes the `*.asn1db` files for all included modules. If they are located in a different directory than the current working directory, use the `include` option to add paths. This is only needed when automatically generating values. For static values using `Value` no options are needed.

asn1rt

Erlang module

Warning:

All functions in this module are deprecated and will be removed in a future release.

Exports

decode(Module,Type,Bytes) -> {ok,Value} | {error,Reason}

Types:

```
Module = Type = atom()  
Value = Reason = term()  
Bytes = binary
```

Decodes Type from Module from the binary Bytes. Returns {ok,Value} if successful.

Use Module:decode(Type, Bytes) instead of this function.

encode(Module,Type,Value)-> {ok,Bytes} | {error,Reason}

Types:

```
Module = Type = atom()  
Value = term()  
Bytes = binary  
Reason = term()
```

Encodes Value of Type defined in the ASN.1 module Module. Returns a binary if successful. To get as fast execution as possible the encode function only performs rudimentary tests that the input Value is a correct instance of Type. The length of strings is, for example, not always checked.

Use Module:encode(Type, Value) instead of this function.

info(Module) -> {ok,Info} | {error,Reason}

Types:

```
Module = atom()  
Info = list()  
Reason = term()
```

info/1 returns the version of the asn1 compiler that was used to compile the module. It also returns the compiler options that was used.

Use Module:info() instead of this function.

utf8_binary_to_list(UTF8Binary) -> {ok,UnicodeList} | {error,Reason}

Types:

```
UTF8Binary = binary()
```

```
UnicodeList = [integer()]  
Reason = term()
```

`utf8_binary_to_list/1` Transforms a UTF8 encoded binary to a list of integers, where each integer represents one character as its unicode value. The function fails if the binary is not a properly encoded UTF8 string.

Use *unicode:characters_to_list/1* instead of this function.

```
utf8_list_to_binary(UnicodeList) -> {ok,UTF8Binary} | {error,Reason}
```

Types:

```
UnicodeList = [integer()]  
UTF8Binary = binary()  
Reason = term()
```

`utf8_list_to_binary/1` Transforms a list of integers, where each integer represents one character as its unicode value, to a UTF8 encoded binary.

Use *unicode:characters_to_binary/1* instead of this function.