The IDP system reference manual

KULeuven Knowledge Representation and Reasoning research group

February 29, 2012

1 Installing And Running

The system has been verified to run under Windows (7), OSX (Lion) and various unix versions.

1.1 Building from source

Required software packages:

- C and C++ compiler, supporting most of the C++11 standard. Examples are GCC 4.4 or higher, clang 3.2 or visual studio 11.
- Cmake build environment.
- Bison and flex packages or yacc and lex packages.
- Pdflatex and doxygen for building the documentation.
- (optional) Gecode for constraint programming support.

Assume idp is unpacked in idpdir, you want to build in builddir (cannot be the same as idpdir) and install in installdir. Building and installing is then achieved by executing the following commands:

```
cd <builddir>
cmake <idpdir> -DCMAKE_INSTALL_PREFIX=<installdir> -DCMAKE_BUILD_TYPE="Release"
make -j 4
make check
make install
```

Alternatively, cmake-gui can be used as a graphical way to set cmake options.

1.2 Running the software

One-shot execution of a procedure proc with a set of files files is achieved by running

```
idp -e "proc()" files
```

An interactive interface using the files files is started with

```
idp -i files
```

Afterwards, command help can be requested with the help() command. Auto-completion is available via the tab key, ctrl-r will search the command history.

Currently including additional files during an interactive session is not supported.

2 Comments

Everything between /* and */ is a comment, as well as everything between // and the end of the line. If a comment block starts with /**, but not with /***, then the comment is added as a description to the first thing after that comment block that can have a description. Currently, only procedures can have a description.

3 Include statements

Everywhere in an IDP file, a statement

```
#include "path/to/file"
```

is replaced by the contents of the file path/to/file. A statement

```
#include <filename>
```

is replaced by the contents of the standard library file filename. Currently the following standard library files are available:

mx Contains some useful model expansion procedures.

4 Namespaces

A namespace with name MySpace is declared by

```
namespace MySpace {
    // content of the namespace
}
```

An object with name MyName declared in namespace MySpace can be referred to by MySpace::MyName. Inside MySpace, MyName can simply be referred to by MyName.

A Namespace can contain namespaces, vocabularies, theories, structures, procedures, options and using statements. A using statement is of one of the following forms

```
using namespace MySpace using vocabulary MyVoc
```

where MySpace is the name of a namespace, and MyVoc the name of a vocabulary. Below such a using statement, objects MyObj declared in MySpace, respectively MyVoc, can be referred to by MyObj, instead of MySpace::MyObj, respectively MyVoc::MyObj.

Every object that is declared outside a namespace, is considered to be part of the global namespace. The name of the global namespace is global_namespace. In other words, every IDP file implicitely starts with namespace global_namespace { and ends wit han extra }.

5 Vocabularies

A vocabulary with name MyVoc is declared by

```
vocabulary MyVoc {
    // contents of the vocabulary
}
```

A vocabulary can contain symbol declarations, symbol pointers, and other vocabularies. Symbols are types (sorts), predicate and functions symbols.

5.1 Symbol declarations

A type with name MyType is declared by

```
type MyType
```

When declaring a type, it can be stated that this type is a subtype or supertype of a set of other types. The following declares MyType to be a supbtype of the previously declared types A1 and A2, and a supertype of the previously declared types B1 and B2:

```
type MyType isa A1, A2 contains B1, B2
```

A predicate with name MyPred, arity 3 and types T1,T2,T3 is declared by

```
MyPred(T1,T2,T3)
```

A predicate with arity zero can be declared by MyPred() or MyPred.

A function with name MyFunc, input types T1,T2,T3 and output type T is declared by

```
MyFunc(T1, T2, T3): T
```

A partial function is declared by

```
partial MyFunc(T1,T2,T3):T
```

Constants of type T can be declared by MyConst:T or MyConst():T. Besides functions with an identifier as name, functions of arity two with names +,-,*,/,% and ^ can be declared, as well as unary functions iwth names - and abs.

5.2 Symbol pointers

To include a type, predicate, or function from a previously declared vocabulary V in another vocabulary W, write

```
/* Declaration of vocabulary V*/
vocabulary V {
    //...
    type A
    P(A)
    F(A,A):A
    //...
}

vocabulary W {
    extern type V::A
    extern V::P[A] //also possible: extern V::P/1
```

```
extern V::F[A,A:A] also possible: extern V::F/2:1
```

In the example, explicitly including type A of vocabulary V in W is not needed, since types of included predicates or functions are automatically included themselves. To include the whole vocabulary V in W at once, used

```
vocabulary W {
    extern vocabulary V
}
```

5.3 The standard vocabulary

The global namespace contains a fixed vocabulary std, which is defined as follows:

```
vocabulary std {
   type nat
   type int contains nat
   type float contains int
   type char
   type string contains char
   +(int,int) : int
   -(int,int) : int
   *(int,int) : int
   /(int,int) : int
   %(int,int) : int
   abs(int) : int
   -(int) : int
   +(float,float) : float
   -(float,float) : float
   *(float,float) : float
   /(float,float) : float
   ^(float,float) : float
   abs(float) : float
   -(float) : float
}
```

Every vocabulary implicitly contains all symbols of std. Also, every vocabulary contains for each of its types A the predicates =(A,A), <(A,A), and >(A,A) and the functions MIN:A, MAX:A, SUCC(A):A and PRED(A):A. In every structure, the symbols of std have the following interpretation:

```
nat
                            all natural numbers
                            all integer numbers
int
                            all floating point numbers
float
                            all characters
char
string
                            all strings
                            integer addition
+(int,int) :
-(int,int):
                            integer subtraction
               int
                            integer multiplication
*(int,int) :
               int
/(int,int) :
               int
                            integer division
                            remainder
%(int,int) :
abs(int) : int
                            absolute value
-(int): int
                            unary minus
+(float,float) : float
                            floating point addition
-(float,float) :
                            floating point subtraction
*(float,float) :
                            floating point multiplication
                    float
/(float,float) : float
                            floating point division
^(float,float) : float
                            floating point exponentiation
abs(float) : float
                            absolute value
-(float) : float
                            unary minus
```

The predicate =/2 is always interpreted by equality. The order $<_{dom}$ on domain elements is defined by

- numers are smaller than non-numers;
- strings are smaller than compound domain elements (see bolow for a definitions of a compound domain element);
- $d_1 <_{dom} d_2$ if d_1 and d_2 are numbers and $d_1 < d_2$;
- $d_1 <_{dom} d_2$ if d_1 and d_2 are strings that are nut numbers and d_1 is before d_2 in the lexicographic ordering:
- $d_1 <_{dom} d_2$ is some total order on compound domain elements (which we do not specify).

Every structure contains the following fixed interpretations:

In an IDP-file, you should disambiguate which MAX you want to use. This is done by MAX[:MyType].

6 Theories

A theory with name MyTheory over a vocabulary MyVoc is declared by

```
theory MyTheory : MyVoc {
    // contents of the theory
}
```

A theory contains sentences and inductive definitions.

6.1 Sentences

6.1.1 Terms

Before explaining the syntax for sentences, we need to introduce the concept of a term and a formula. We also give the syntax for terms and formulas in IDP.

A *term* is inductively defined as follows:

- a variable is a term;
- a constant is a term;
- if F is a function symbol with n input arguments and t_1, \ldots, t_n are terms, then $F(t_1, \ldots, t_n)$ is a term.

In IDP, variables start with a letter and may contain letters, digits and underscores. When writing a term in IDP, the constant and function symbols occurring in that term should be declared before. The *type of a term* is defined as its return type (see section 5.1) in the case of constants and functions. The type of a variable is derived from its occurrences in formulas (see section 6.5). If a term occurs in an input position of a function, then the type of the term and the type of the input position must have a common ancestor type.

6.1.2 Formulas and Sentences

A *formula* is inductively defined by:

- true and false are formulas;
- if P is a predicate symbol with arity n and t_1, \ldots, t_n are terms, then $P(t_1, \ldots, t_n)$ is a formula;
- if t_1 and t_2 are terms, then $t_1 = t_2$ is a formula;
- if φ and ψ are formulas and x is a variable, then the following are formulas: $\neg \varphi$, $\varphi \land \psi$, $\varphi \lor \psi$, $\varphi \Rightarrow \psi$, $\varphi \Leftarrow \psi$, $\varphi \equiv \psi$, $\forall x \varphi$, and $\exists x \varphi$.

The following order of binding is used: \neg binds tightest, next \land and \lor , then \Rightarrow and \equiv , and finally \forall and \exists . Desambiguation can be done using brackets '(' and ')'. E.g. the formula $\forall x \ P(x) \land \neg Q(x) \Rightarrow R(x)$ is equivalent to the formula $\forall x \ ((P(x) \land (\neg Q(x))) \Rightarrow R(x))$.

As for terms, if term t occurs in predicate P, then the type of t and the type of the input position of P where it occurs must have a common ancestor type. For formulas of the form $t_1 = t_2$, t_1 and t_2 must have a common ancestor type.

The scope of a quantification $\forall x$ or $\exists x$, is the quantified formula. E.g., in $\forall x \ \psi$, the scope of $\forall x$ is the formula ψ . An occurrence of a variable x that is not inside the scope of a quantification $\forall x$ or $\exists x$ is called *free*. A sentence is a formula containing no free occurrences of variables. If an IDP problem specification contains formulas that are not sentences, the system will implicitly quantify this variable universally and return a warning message, specifying which variables occur free. Each sentence in IDP should end with a dot '.'.

The IDP syntax of the different symbols in formulas are given in the table below. Also the informal meaning of the symbols is given.

Logic	IDP	Declarative reading
\wedge	&	and
\vee		or
\neg	~	not
\Rightarrow	=>	implies
\Leftarrow	=> <= <=>	is implied by
=	<=>	is equivalent to
\forall	!	for each
3	?	there exists
=	=	equals
\neq	~=	does not equal

Besides this, for every natural number n, IDP also supports the following quantifiers (with their respective meanings):

IDP	Declarative reading		
?n	there exist n different elements such that		
? <n< td=""><td colspan="2">there exist less than n</td></n<>	there exist less than n		
?= <n< td=""><td colspan="2">there exist at most n</td></n<>	there exist at most n		
?=n	there exist exactly n		
?>n	there exist more than n		

A universally quantified formula $\forall x \ P(x)$ becomes '! x : P(x)' in IDP syntax, and similarly for existentially quantified formulas. As a shorthand for the formula '! x : ! y : ! z : Q(x,y,z).', one can write '! x y z : Q(x,y,z)'.

In IDP, every variable has a type. The informal meaning of a sentence of the form $\forall x \ \psi$, respectively $\exists x \ \psi$, where x has type T is then 'for each object x of type T, ψ must be true', respectively 'there exists at least one object x of type T such that ψ is true'. The type of a variable can be declared by the user, or derived by IDP (see section 6.5).

6.1.3 Definitions

A definition defines a concept, i.e. a predicate, in terms of other predicates. Formally, a definition is a set of rules of the form

$$\forall x_1, \dots, x_n \ P(t_1, \dots, t_m) \leftarrow \varphi$$

where P is a predicate symbol, t_1, \ldots, t_m are terms that may contain the variables x_1, \ldots, x_n and φ a formula that may contain these variables. $P(t_1, \ldots, t_m)$ is called the *head* of the rule and ψ the *body*.

A definition in IDP syntax consists of a set of rules, enclosed by '{' and '}'. Each rule ends with a '.'. The definitional implication \leftarrow is written '<-'. The quantifications before the head may be omitted in IDP, i.e., all free variables of a rule are implicitly universally quantified. If the body of a rule is empty, the rule symbol '<-' can be omitted. Recursive definitions are allowed in IDP. The semantics for a definitions are the wellfounded semantics TODO: reference.

6.2 Chains of (in)equalities

As in mathematics, one can write chains of (in)equalities in IDP. They can be used as shorthands for conjunctions of (in)equalities. E.g.:

```
! x y : (1 = < x < y = < 5) => ...

// is a shorthand for

! x y : (1 = < x) & (x < y) & (y = < 5) => ...
```

6.3 Aggregates

Aggregates are functions that take a set as argument, instead of a simple variable. IDP supports some aggregates that map a set to an integer. As such, they can be seen as integer terms. There are two kinds of sets in IDP.

- An expression of the form '[(phi_1,t_1) ; (phi_2,t_2) ; ... ; (phi_n,t_n)]', where each phi_i is a formula and each t_i is a term.
- An expression of the form ' $\{x_1 x_2 \dots x_n : phi:t\}$ ', where the x_i are variables, phi is a formula and t is a term.

The current system has support for five aggregate functions:

Cardinality: The cardinality of a set is the number of elements in that set. The IDP syntax for the cardinality of a set S is 'card S' or '# S'. For the first kind of sets, this denotes the number of formulas phi_i that are true. For the second kind, this is interpreted as the number of tuples (a_1,a_2,..., a_n) such that phi is true.

Sum: Let S be a set of the second form, i.e., of the form '{ $x_1 x_2 \dots x_n : phi: t$ }'. Then the interpretation of 'sum S' denotes the number

$$\sum_{(\mathtt{a_1},\mathtt{a_2},\ldots,\mathtt{a_n})|I\vDash \mathtt{phi}}\mathtt{t},$$

i.e., it is the sum of all the terms for which there exist a_1,..., a_n that make the formula phi true. For sets of the first sort, this is interpreted as

$$\sum_{i|I\models \mathtt{phi}_\mathtt{i}} \mathtt{t}_\mathtt{i}.$$

Product: Products are defined similar to sum.

Maximum: One can write 'max S' to denote the maximum value of the term in S, i.e.,

$$\max_{(\mathtt{a_1},\mathtt{a_2},\ldots,\mathtt{a_n})|I\vDash \mathtt{phi}}\mathtt{t}$$

for sets of the second sort. Sets of the first sort are handled analogously.

Minimum: To get the minimum value, write 'min S'.

When using cardinality, the terms do not matter. You can choose to write 1 for every term, but are also allowed to leave out the terms.

6.4 Partial functions

A normal function is total: it assigns an output value to each of its input values. On the other hand, partial functions do not necessarily have this property. In IDP, partial function F can arise in different situations. Either F is explicitly declared as partial function, or it is declared total, but its input types or output type are subtypes or integer types.

The semantics of a partial function F is given by transforming constraints and rules where F occurs as follows:

- in a positive context, $P(\ldots, F(x), \ldots)$ is transformed to $\forall y \ (F(x) = y \Rightarrow P(\ldots, y, \ldots);$
- in a negative context, $P(\ldots, F(x), \ldots)$ is transformed to $\exists y \ (F(y) = y \land P(\ldots, y, \ldots)$.

Here, P(..., F(x),...) occurs in a positive context if it occurs in sentence and in the scope of an even number of negations, or it occurs in a body of a rule and in the scope of an odd number of negations. All other occurrences are in a negative context.

6.5 The Type of a Variable

There are two ways to assign a type t to a variable v:

• Explicitly mention the type of v between '[' and ']' when v is quantified. Then v gets type t in the scope of the quantifier. E.g.,

```
theory T: V {
   ! MyVar[MyType] : ? MyVar2[MyType2] MyVar3[MyType3] : //...
}
```

• Do not mention the type of v but let the system automatically derive it. The rest of this section explains how this is done.

6.5.1 Automatic derivation of types for variables

We distinguish between typed and untyped occurrences. The following are typed occurrences of a variable x:

- an occurrence as argument of a non-overloaded predicate: $P(\ldots, x, \ldots)$;
- an occurrence as argument of a non-overloaded function: $F(\ldots,x,\ldots)=\ldots$;

• an occurrence as return value of a non-overloaded function: $F(\ldots) = x$ or $F(\ldots) \neq x$.

All others positions are untyped.

An overloaded predicate or function symbol can be disambiguated by scecifying its vocabluary and / or types. E.g.,

```
! x: MyVoc::P[A,A](x,x).
! y: ?1 x : F[A:A](x) = y.
MyVoc::C[:A] > 2.
```

In this case, te occurences of all variables are typed.

Basically, if a variable occurs in a typed position, it gets the type of that position. If a declared variable with type T_1 occurs in a typed position of type T_2, then T_1 and T_2 should have a common ancestor type.

The more complicated cases arise when a variable does not occur in any typed position, or it occurs in two typed positions with a different type. The system is designed to give a reasonable type to such variables. However, the choices made by the system are ad hoc and are probably not the ones the user intended. TODO: stukje over wat typederivation precies doet voor gelijkheid en zo.. Broes?

First consider the case where a variable occurs in typed positions with different types. The IDP system will then give a warning. If all the typed positions where the variable occurs have a common ancestor type T, then the variable is assigned this type T. If they do not have a common ancestor, no derivation is done.

Now consider the case where a variable does not occur in a typed position. Then, the IDP system tries to find out what the type of the variable should be using its occurences in untyped position in built-in overloaded functions. For example, when a variable x only occurs in x = t, then x will get the same type as t. This behaviour might not always be the desired, so the IDP system will give a warning, including which type it derived for the variable. It's always safer to declare a type for the variable in this case. If it is not possible to derive a type for x in this way either, the IDP system reports an error.

7 Structures

A (three-valued) structure with name MyStruct over a vocabulary MyVoc is declared by

```
structure MyStruct: MyVoc {
    //contents of the structures
}
or by
asp_structure MyStruct: MyVoc {
    //contents of the structures
}
```

7.1 Contents of a structure

A particular input to a problem can be given by giving a (three valued) interpretation to all types and some predicate and function symbols of a given vocabulary. Here, we describe the different ways to specify a structure.

7.1.1 Type Enumeration

The syntax for a type enumeration is

```
MyType = { El_1; El_2; ...; El_n }
```

where MyType is the name of the enumerated type and El_1; El_2; ...; El_n are the names of the objects of that type. Names of objects can be (positive and negative) integers, strings, chars, compound domain elements, or identifiers that start with an upper- or lowercase letter. Identifiers are shorthands for strings (without the quotes) and can be interchanged. TODO: Important: in the lua code, the identifiers cannot be used safely, only their string equivalent If one type is a subtype of another, all elements of the subtype are added to the supertype also. In the case all subtypes of a given type are specified, the supertype is derived to be the union of all elements of the subtypes. If a type is not specified, all domain elements of that type that occur in a predicate or function interpretation (see below) are automatically added to that type.

7.1.2 Predicate Enumeration

The syntax for enumerating all tuples for which a predicate MyPred with n arguments is true is as follows.

It is also possible to write parentheses around tuples.

This notation makes it possible to state that a proposition (a predicate with no arguments) is true, by using an empty tuple.

```
true = { () }
false = { }
```

However, it might be easier to use true and false instead of { () } and {}.

7.1.3 Function Enumeration

The syntax for enumerating a function MyFunc with n arguments is

To give the interpretation of a constant, one can simply write 'MyConst = El' instead of 'MyConst = { -> El }'.

7.1.4 Compound Domain Elements

A function applied to a tuple of domain elements can be used as a domain element. We call such a domain element a *compound domain element*. An example is the domain element F(1,a). If F/n is a function then

```
F = generate
```

specifies that the interpretation of F is the two-valued interpretation that maps each tuple (d_1, \ldots, d_n) to the compound domain element $F(d_1, \ldots, d_n)$.

7.1.5 Three-Valued Predicate/Function interpretations

Three-valued interpretations are given by either

- enumerating the certainly true and certainly false tuples;
- enumerating the certainly true and the unknown tuples;
- enumerating the unknown and the certainly false tuples.

To sepcify which tuples are enumerated, use <ct>, <cf> and <u>. For example

```
P < ct > = \{ /* \text{ enumeration of the certainly true tuples of P */ } 

P < u > = \{ /* \text{ enumeration of the unknown tuples of P */ }
```

7.1.6 Interpretation by Procedures

The syntax

```
P = procedure MyProc
```

is used to interpret a predicate or function symbol P by a procedure MyProc (see below). If P is an n-ary predicate, then MyProc should be an n-ary procedure that returns a boolean. If P is an n-ary function, then MyProc should be and n-ary function that returns a number, string, or compound domain element.

7.1.7 Shorthands

Shorthands like 'MyType = {1..10; 15..20}' or 'MyType = { a..e; A..E }' may be used for enumerating types or predicates with only one argument.

7.2 ASP structures

An ASP structure consists of a list of facts in the usual ASP syntax. In particular, everything from a % till the end of the line is considered a comment, and - before an atom denotes classical negation (negation as failure is not available). A fact about functions is written like F(a)=b or -F(c)=d.

8 Procedures

8.1 Declaring a procedure

A procedure with name MyProc and arguments A1, ..., An is declared by

```
procedure MyProc(A1,...,An) {
    // contents of the procedure
}
```

Inside a procedure, any chunk of Lua code can be written. For Lua's reference manual, see http://www.lua.org/manual/5.1/. In the following, we assume that the reader is familiar with the basic concepts of Lua.

8.2 IDP types

Besides the standard types of variables available in Lua, the following extra types are available in IDP procedures.

sort A set of sorts with the same name. Can be used as a single sort if the set is a singleton.

- **predicate_symbol** A set of predicates with the same name, but possibly with different arities. Can be used as a single predicate if the set is a singleton. If P is a predicate_symbol and n an integer, then P/n returns a predicate_symbol containing all predicates in P with arity n. If $s1, \ldots, sn$ are sorts, then P[$s1, \ldots, sn$] returns a predicate_symbol containing all predicates Q/n in P, such that the *i*'th sort of Q belongs to the set Si, for $1 \le i \le n$.
- function_symbol A set of first-order functions with the same name, but possibly with different arities. Can be used as a single first-order function if the set is a singleton. If F is a function_symbol and n an integer, then F/n:1 returns a function_symbol containing all function in F with arity n. If $s1, \ldots, sn$, t are sorts, then $F[s1, \ldots, sn:t]$ returns a function_symbol containing all functions G/n in F, such that the i'th sort of F belongs to the set si, for $1 \le i \le n$, and the output sort of G belongs to t.
- **symbol** A set of symbols of a vocabulary with the same name. Can be used as if it were a sort, predicate_symbol, or function_symbol.
- vocabulary A vocabulary. If V is a vocabulary and s a string, V[s] returns the symbols in V with name s.
- **compound** A domain element of the form $F(d_1, \ldots, d_n)$, where F is a first-order function and d_1, \ldots, d_n are domain elements.
- tuple A tuple of domain elements. T[n] returns the n'th element in tuple T.

predicate_table A table of tuples of domain elements.

function_interpretation An interpretation for a function. F.graph returns the predicate_interpretation of the graph associated to the function_interpretation F.

structure A first-order structure. To obtain the interpretation of a sort, singleton predicate_symbol, or singleton function_symbol symb in structure S, write S[symb].

theory A logic theory.

options A set of options.

namespace A namespace.

overloaded An overloaded object.

8.3 Built-in procedures

A lot of procedures are allready built in. Typing help in interactive mode shows an overview of the available procedures, together with a description.

9 Options

The IDP system has various options. To set an option, you can use the following lua-code

```
stdoptions.MyOption = MyValue
```

where MyOption is the name of the option and MyValue is the value you want to give it. If you want to have multiple option sets, you can make them with them with

```
FirstOptionSet = newOptions()
SecondOptionSet = newOptions()
FirstOptionSet.MyOption = MyValue
SecondOptionSet.MyOption = MyValue
```

To activate an option set, use the procedure setascurrentoptions(MyOptionSet). From that moment, MyOptionSet will be used in all comands.

autocomplete = [false, true] Turn autocompletion of structures on or off

groundverbosity = [0..max(int)] Verbosity of the grounder. The higher the verbosity, the more debug information is printed.

language = [ecnf, idp, tptp] The language used when printing objects.

longnames = [false, true] If true, everything is printed with reference to their vocabulary. For
 example, a predicate P from vocabulary V will be printed as V::P instead of P.

nbmodels = [0..max(int)] Set the number of models wanted from the model expansion inference. If set to 0, all models are returned.

satverbosity = [0..max(int)] Like groundverbosity, but controls the verbosity of MiniSAT(ID)

timeout = [0..max(int)] Set the timeout for inferences (in seconds)

 ${\bf trace}=[{\bf false},{\bf true}]$ If true, the procedure model expand produces also an execution trace of MiniSAT(ID)