



---

**IST FP6-004779**

**PYPY**

**Researching a Highly Flexible and Modular Language Platform and Implementing it  
by Leveraging the Open Source Python Language and Community**

**STREP**

**IST Priority 2**

## **D9.1 Constraint Satisfaction and Inference Engine, Logic Programming in Python**

**Due date of deliverable: December 31, 2006**

**Actual Submission date: May 11, 2007**

**Start date of Project: 1st December 2004**

**Duration: 28 months**

**Lead Contractor of this WP: DFKI**

**Authors: Aurélien Campéas (Logilab), Anders Lehmann, Stephan Busemann (DFKI)**

**Revision: Final**

**Project co-funded by the European Commission within the Sixth Framework  
Programme (2002-2006)**

**Dissemination Level: PU (Public)**



---

## Revision History

Date	Name	Reason of Change
June 01	Anders Lehmann	Created document, OWL and Semantic Web
July 07	Aurélien Campéas	Logilab contributions: state of the art, logic and constraint programming
July 13	Anders Lehmann	Revisions to OWL part (DFKI)
July 27	Aurélien Campéas	Extensions and revisions (logilab part)
July 28	Stephan Busemann	First published draft, formal matters
Dec 14	Aurélien Campéas	Considering external review
Feb 27	Stephan Busemann	Interim version to be sent to EC
Mar 07	Aurélien Campéas	Famous last words
Mar 13	Anders Lehmann	Added/changed subsections of OWL implementation
May 07	Stephan Busemann	Update OWL part (DFKI), introduction, References
May 11	Carl Friedrich Bolz	Publish Final Version on the web page

## Abstract

This report describes the concurrent logic and constraint programming approaches taken in PyPy. Following some historical background, logic programming in PyPy is presented. Python is extended by the concept of logic variable that is implemented in a special object space, along with a non-deterministic choice construct. Constraint Programming allows one to specify and solve constraint satisfaction problems. Methods to define and extend solvers are described. A reasoning module is also described, using the semantic web language OWL. An application of OWL-based reasoning to human-language question answering is described.



## Contents

<b>1</b>	<b>Executive Summary</b>	<b>5</b>
<b>2</b>	<b>Introduction</b>	<b>5</b>
2.1	Purpose of this Document . . . . .	5
2.2	Scope of this Document . . . . .	5
2.3	Related Documents . . . . .	6
2.4	Getting Started . . . . .	6
<b>3</b>	<b>Concurrent Constraint and Logic Programming</b>	<b>6</b>
3.1	Motivation . . . . .	6
3.2	Logic Programming in Prolog . . . . .	7
3.3	Constraint Programming and MAC . . . . .	7
3.4	Limits of Prolog and MAC . . . . .	7
3.5	From Backtracking Engines to CCLP . . . . .	8
3.6	From CCLP to Oz to Python (anecdotal bits) . . . . .	8
3.7	Constraint Handling Rules . . . . .	9
<b>4</b>	<b>Logic Programming in PyPy</b>	<b>9</b>
4.1	Basic Elements Serving Deterministic Logic Programming . . . . .	9
4.2	Threads and Dataflow Synchronization . . . . .	11
4.3	Table of Operators . . . . .	12
4.4	Implementation Aspects . . . . .	13
<b>5</b>	<b>Programming With Search</b>	<b>13</b>
5.1	Non-Deterministic Logic Programs . . . . .	13
5.2	Constraint programs . . . . .	16
5.3	Using the constraint engine . . . . .	16
5.4	Table of Operators . . . . .	17
5.5	Extending the Search Engine . . . . .	17
5.6	Implementation Aspects . . . . .	19
5.7	History of the Implementation, Credits . . . . .	19
<b>6</b>	<b>Semantic Web: Reasoning in OWL</b>	<b>20</b>
6.1	Designing an OWL-DL Reasoner . . . . .	20
6.2	Implementation of Builtin OWL Types . . . . .	21
6.3	Implementation of Builtin OWL Predicates . . . . .	22
6.4	Example of Using the OWL Reasoner . . . . .	22



---

<b>7</b>	<b>Using the Results for Semantic Web Applications</b>	<b>23</b>
7.1	Querying the Language Technology World . . . . .	23
7.2	Using SPARQL as a query language . . . . .	24
7.3	An interpreter for query answering . . . . .	25
7.4	Performance issues . . . . .	25
7.5	Installation details . . . . .	26
7.6	State of implementation . . . . .	26
<b>8</b>	<b>Conclusion</b>	<b>26</b>
<b>9</b>	<b>Glossary of Abbreviations</b>	<b>27</b>
9.1	Technical Abbreviations: . . . . .	27
9.2	Partner Acronyms: . . . . .	28



## 1 Executive Summary

This report discusses logic programming and constraint programming over discrete domains in PyPy with an application to an OWL reasoner. Most of the features considered thereafter are implemented in a special object space. This demonstrates the flexibility of the PyPy infrastructure, the coroutine machinery and parts of the garbage collector framework, as described in [D7.1]. The OWL reasoner uses the constraint solver for making inferences on OWL ontologies.

The distinctive features added to PyPy in this work are subsumed under the concept of Concurrent Constraint and Logic Programming (abbreviated as CCLP in the remainder of this document).

Several code snippets are presented, along with discussion on the state of the implementation of the various sub-components involved.

## 2 Introduction

### 2.1 Purpose of this Document

This report describes the concurrent logic and constraint programming approaches taken in PyPy that were used to implement an OWL reasoner. The reasoner was equipped with a SPARQL interface and applied in the context of query answering.

### 2.2 Scope of this Document

This document describes the concept and implementation of a framework for concurrent constraint and logic programming in Python.

DFKI and Logilab started with an exploratory implementation of a constraint solver using the Computation Space abstraction. Also logic variables, along with unification and dataflow thread synchronization were experimented with (using Python threads and condition variables).

The Logic Object Space was implemented by Strakt, Merlinux, Logilab and DFKI. This work has been completed for all deterministic logic operators, a specific interpreter-level scheduler, exception propagation semantics between threads bound by logic variables, and preliminary support of non-deterministic computation. The initial pure-python constraint solver has been partly ported to RPython.

There is on going work on the implementation and test of the computation space abstraction.

The OWL reasoner was implemented by DFKI together with an interface to the query language SPARQL [SPQL]. An application to the ontology of the “world of language technology” (<http://www.lt-world.org>) was realized that relies on SPARQL.

While this application was foreseen to be within the scope of this document, it could not finally be tested and evaluated at the time this report is submitted. The delay is due to bugs found in the OWL resolver. DFKI is supported by its former key developer, Anders Lehmann, who had sought other challenges in December 2006. Testing and evaluating the application is the only part considered non-final. The sections on the application have the details. We expect to present a relevant demonstration by May 31st, 2007.

We thank Marian Babik, Institute of Informatics, Slovak Academy of Sciences, who kindly agreed to review an intermediate version of this document as an external reviewer. His comments and recommendations, which were received on Dec 1, 2006, proved very valuable.



## 2.3 Related Documents

Support for Massive Parallelism and Publish about Optimization results, Practical Usages and Approaches for Translation Aspects [D7.1].

## 2.4 Getting Started

To try the examples in this document one should compile and run a standalone interpreter:

```
cd pypy/translator/goal
python2.4 ./translate.py targetpypystandalone.py -o logic
```

Alas, not all examples will work, due to the state of the implementation. Operations related to search with logic programs don't work. See the *history of the implementation* section for details. However, the non-concurrent version of the constraint solver is usable.

## 3 Concurrent Constraint and Logic Programming

### 3.1 Motivation

The first step of this work was to research modern logic and constraint programming systems supporting concurrency.

The motivation was to add to Python a set of primitives enabling logic and constraint programming. Indeed, using an external library written in an efficient, compiled language is the only option currently available to the Python programmer, if she wants raw performance. And, to our knowledge, no binding for such libraries (such as [gecode](#) or [choco](#)) exists today.

On the other hand, some pure Python solutions can be found. Notably, [logilab's constraint package](#) (which implements Oz's algorithms [ECP94]) and [Gustavo Niemeyer's constraint solver](#) (implementing MAC [MAC97]) are examples. We know for sure that logilab's package is used daily in production. However, from personal experience it is known that for some moderately sized problems, performance is inadequate.

People with whom we discussed about the availability of logic or constraint tools in Python regularly mentioned other tools, such as Pylog<sup>1</sup>, Metalog<sup>2</sup>, Psychinko<sup>3</sup>, CWM<sup>4</sup>. The first (and only one concerned with logic programming) is quite unsatisfying in that there can be only a limited set of data structures common to Python and SWI-Prolog. Metalog is biased towards semantic web applications, and also since it is written in pure Python, its engine can run into the performance problems of the aforementioned constraint packages. The others are concerned with production rule systems, which while useful in their own right and also in conjunction with logic/constraint systems, do not provide the same services.

There was clearly some unsatisfied demand there. Thus it made sense to incorporate into PyPy one solution that would comply with many non-obvious requirements:

1. build logic and constraint programming on top of some common infrastructure,

---

<sup>1</sup>Pylog is a bridge linking CPython to SWI Prolog.

<sup>2</sup>Metalog "is a next-generation reasoning system for the Semantic Web. Historically, Metalog has been the first semantic web system to be designed, introducing reasoning within the Semantic Web infrastructure by adding the query/logical layer on top of RDF."

<sup>3</sup>Psychinko is a Python implementation of the RETE algorithm, aimed at production systems. RETE is quite different from logic/constraint programming as it is concerned with truth maintenance in the face of a changing environment.

<sup>4</sup>CWM is "a popular Semantic Web program that can, amongst many things, perform inferences as a forward chaining FOPL inference engine". CWM uses Psychinko.



2. have the logic/constraint programming styles be maximally integrated with the imperative object-oriented language that is Python,
3. support concurrency (because interactive programs need it), multi-processing and distribution (because exploration of vast search spaces is a processing power hungry activity),
4. be flexible and extensible,
5. be orders of magnitude faster than a pure Python implementation.

This might well be completely impossible within the current CPython implementation, which many developers deem inflexible and hard to adapt. We still doubt it could be done like we did with other Python implementations, like Jython or IronPython. Indeed, we exploited some features quite unique to PyPy in order to achieve a solution which satisfies all these constraints. This is all explained in the rest of this report.

Prolog [[PRLG92](#)] and constraint solving algorithms like MAC [[MAC97](#)] were considered, but we finally decided to go with the more recent paradigm named Concurrent Constraint and Logic Programming [[CCLP93](#)], as it supports requirement #1, #2 and #3. More precisely, we decided to walk in the steps of the [Oz](#) programming language [[OPM95](#)], which provides a CCLP framework also satisfying #4. Leveraging the PyPy architecture allowed us to implement all of this while respecting #5.

We now skim very briefly over the history and state of affairs in the realm of logic and constraint programming, in order to expose the rationale behind our architectural choices.

## 3.2 Logic Programming in Prolog

While first touched by John McCarthy in *Programs with common sense* [[MCC59](#)], and approached in languages like [Planner](#), logic programming was installed firmly with Prolog. Despite Prolog having incomplete logical semantics (due to its depth-first search strategy and the existence of cut operator), it has found a vast range of successful applications -- from (natural) language processing to sophisticated user interface development, passing by knowledge management and expert systems.

The algorithmic basis of Prolog is [first-order resolution](#), which combines [backtracking](#) and [unification](#) of variables.

## 3.3 Constraint Programming and MAC

The MAC (maintain arc consistency) algorithm for solving constraint satisfaction problems is currently considered the best in its category. One important vendor, [ILOG](#), based its commercial solver on it with great success.

The algorithmic bases of MAC are [backtracking](#) and *arc-consistency checking* [[AIJ94](#)] interleaved.

## 3.4 Limits of Prolog and MAC

A typical classic Prolog implementation and MAC share several fundamental common mechanisms (backtracking and trailing); as such, modern Prolog implementations (such as [GNU Prolog](#)) have been augmented with finite domain constraint programming without much pain.

Both Prolog and MAC, however can be considered lacking [[PVR02](#)] with respect to the following aspects:

- the search strategy is hard-wired into the engine; a different search strategy would have to be implemented, not easily, at the program level;



- it is hard to make them efficiently concurrent, which is due to the use of trailing (stack disciplined, and data-type dependant memorization of previous bindings) as technique to remember the previous values taken before engaging in a choice point (they can be distributed, but without exploiting the inherent parallelism or concurrency of a distributed environment);

Specific to Prolog, as a programming language:

- it tries to satisfy both the need for search and algorithmic problems (those that have known efficient algorithms); the resulting compromise is insufficient on both sides;
- it does not support higher-order programming; it is not possible to mix functional style, logic style and even imperative style;

Constraints in MAC solver are not restricted to [Horn Clause](#) style; their expression is bound to whatever language the implementor uses (in the case of ILOG solver, for instance, it is C++).

### 3.5 From Backtracking Engines to CCLP

In short, CCLP draws from research in parallelization and modularization of logic programming, integration with stateful computing and integration with constraint programming. CCLP emerged after almost two decades of research into ways to extract the anticipated inherent parallelism of a logic language.

Concurrent logic languages had a rich history with experiments like *Parlog* [[PAR86](#)], *Concurrent Prolog* [[CPR87](#)], *Guarded Horn Clauses* [[GHC88](#)].

The most significant step was the *ask* and *tell* operators brought by *Saraswat* [[CCP91](#)]. This work reconciliated parallel logic languages with constraint programming. The Andorra Kernel Language [[AKL91](#)], that embodied these ideas, added state (with ports), and encapsulated search.

It appeared that in CCLP, the logic or constraint based mechanism could be seen as a concurrent process calculus. *Ask* is indeed a synchronization primitive, and shared variables serve as communication devices between processes. Concurrent programming in this style alleviates the imperative style riddled with low-level and error-prone locks, mutexes and semaphores. It also has clean logic semantics.

### 3.6 From CCLP to Oz to Python (anecdotal bits)

Expanding on AKL, the [Oz](#) language integrates CCLP into an already functional and also stateful programming framework, which is truly multi-paradigm. In this light, it was decided to meet the current Oz developer group in Louvain-la-Neuve to get insights on the feasibility of extending Python with the powerful abstractions present in Oz. We had to understand how one single language can bring so many seemingly different techniques into a coherent whole.

The meeting was fruitful and quite enthusiastic (some people working on Oz happen to be happy Python users). Some skepticism was shed with respect to Python's informal, and in fact very complex semantics, and how difficult it could be to add something radically different like logic programming as first class citizen in such a language. It was stated that the clean, formalized and well-understood, onion-structured semantic layers of Oz was one key to the harmony of the whole.

More seriously, the encapsulated search principle deployed with AKL has been shown to pave the road to languages mixing stateful and logic/constraint programming. The ability to copy whole computations (where a computation is defined as a bag of data comprising a set of threads plus its associated heap) eliminates the difficulty of implementing custom trailing semantics for custom ad hoc data types (or objects). It does not completely alleviate the dangers inherent to the manipulation of mutable data shared accross encapsulated computations (thus





obviously violating encapsulation) though. But one usual answer to such concerns is that programmer is responsible for the proper encapsulation of different styles in a modular way. With power comes responsibility. This bodes well with the Python philosophy of considering the programmer a responsible grown-up (most of the time).

We do not know yet, of course, if our extension to Python will prove practical in the long run; it could fail either because of language issues (we can see right now that the non-declaration of variables in Python is a bit of a concern) or because these styles are not well understood by most programmers.

The most important (and novel) concept imported from Oz is that of Computation Space. We will not, in the interim report, define it formally, since it has been done in other places [PCS00]; it is nonetheless exposed in the sections that deal with logic programming and constraint programming in PyPy.

## 3.7 Constraint Handling Rules

Constraint Handling Rules [CHR98] are a special purpose constraint definition language for user-defined constraints. They complement CCLP languages. They bring features such as automatic simplification of constraints (preserving logical equivalence) and propagation (adding redundant equivalent constraints, which may cause further simplification) over these. We don't have this in the context of the WP09 work. It is a bit too early to decide if, when and how to achieve that.

It must be noted that the need for Constraint Handling Rules can be mitigated by the availability of first class constraints, and also the advanced computation space features described in [PCS00]. Neither feature will be provided as part of the WP09 work.

Operations on constraints can always be retrofitted later into the framework.

## 4 Logic Programming in PyPy

The first addition to standard Python is the concept of logic variable. Logic variables serve two main, distinct purposes:

- they are a building block for a declarative style of programming, in which a program has obvious logical semantics (by contrast with a random Python program, whose logical semantics can become obscure extremely quickly);
- they dynamically drive the scheduling of threads in a multi-threaded system, allowing lock-free<sup>5</sup> concurrent programming.

In this section we describe the design, usage and implementation of logic variables in PyPy.

### 4.1 Basic Elements Serving Deterministic Logic Programming

Logic variables are similar to Python variables in the following sense: they map names to values in a defined scope. But unlike normal Python variables, they have *only two states*: free and bound. A bound logic variable is indistinguishable from a normal Python value, which it wraps. A free variable can only be bound once (it is also said to be a single-assignment variable). Due to the extended PyPy still using Python variables, it is good practice to denote logic variables with a beginning capital letter, so as to avoid confusion.

The following snippet creates a new logic variable and asserts its state:

---

<sup>5</sup>lock-free but not block-free, one symptom of a badly written concurrent program using logic variables being that it hangs indefinitely.



```
X = newvar()
assert is_free(X)
```

Logic variables are bound thusly:

```
bind(X, 42)
assert is_bound(X)
assert X / 2 == 21
```

The single-assignment property is easily checked:

```
bind(X, 'hello') # would raise a FailureException
bind(X, 42)      # is tolerated (it does not break monotonicity)
```

It should be quite obvious from this that logic variables are really objects acting as boxes for Python values. No syntactic extension to Python is provided yet to lessen this inconvenience.

The bind operator is low-level. The more general operation that binds a logic variable is known as [unification](#). Unify is an operator that takes two arbitrary data structures and tries to assert their equalness, much in the sense of the Python `__eq__` operator, but with one important twist: unify mutates the state of the involved logic variables.

The following basic example shows logic variables embedded into dictionaries:

```
Z, W = newvar(), newvar()
unify({'a': 42, 'b': Z},
      {'a': Z, 'b': W})
assert Z == W == 42
```

Unify is thus defined as follows (it is symmetric):

Unify	value	unbound var
value	equal?	bind
unbound var	bind	alias

Given two values, it does nothing if they are equal, or raises a `UnificationError` exception. Binding happens whenever one value is unified with an unbound variable. Finally, two unbound variables are aliased, that is, constrained to be bound by the same future value.

We now turn to an example involving custom data types:

```
class Foo(object):
    def __init__(self, a):
        self.a = a
        self.b = newvar()

f1 = Foo(newvar())
f2 = Foo(42)
unify(f1, f2)
assert f1.a == f2.a == 42    # assert (a)
assert alias_of(f1.b, f2.b)  # assert (b)
unify(f2.b, 'foo')
assert f1.b == f2.b == 'foo' # (b) is entailed indeed
```



Finally, note that by stating:

```
X = newvar()
X = 42
```

the logic variable referred to by `X` is actually replaced by the value 42. This is not equivalent to `bind(X, 42)`.

What is provided here is sufficient to achieve deterministic logic programming: we need an additional operator to declare non-deterministic choice points in order to get the expressive power of [pure Prolog](#) (understood as Prolog without cut, and with complete logical semantics).

## 4.2 Threads and Dataflow Synchronization

When a thread tries to access a free logic variable, it is suspended until the variable is bound. This behaviour is known as “dataflow synchronization”. With respect to behaviour under concurrency conditions, logic variables come with two operators:

- `wait/1` (defined on a variable) suspends the current thread until the variable is bound, it returns the value otherwise,
- `wait_needed/1` (on a variable also) suspends the current thread until the variable has received a wait message. It has to be used explicitly, typically by a producer thread that wants to lazily produce data.

In this context, binding a variable to a value will make runnable all threads previously blocked on this variable.

Using the “future” builtin, which spawns a coroutine and applies the second to `nth` arguments to its first argument, we show how to implement a producer/consumer scheme:

```
def generate(n, limit):
    if n < limit:
        return (n, generate(n + 1, limit))
    return None

def sum(L, a):
    Head, Tail = newvar(), newvar()
    unify(L, (Head, Tail))
    if Tail != None:
        return sum(Tail, Head + a)
    return a + Head

X = newvar()
S = newvar()

from cclp import stacklet
unify(S, future(sum, X, 0))
unify(X, future(generate, 0, 10))

assert S == 45
```

Note that this eagerly generates all elements before the first of them is consumed. `Wait_needed` helps us write a lazy version of the generator. But the consumer will be responsible of the termination and must thus be adapted, too:



```
def lgenerate(n, L):
    """lazy version of generate"""
    wait_needed(L)
    Tail = newvar()
    bind(L, (n, Tail))
    lgenerate(n+1, Tail)

def lsum(L, a, limit, R):
    """this summer controls the generator"""
    if limit > 0:
        Head, Tail = newvar(), newvar()
        wait(L)
        unify(L, (Head, Tail))
        return lsum(Tail, a+Head, limit-1, R)
    else:
        bind(R, a)

Y = newvar()
T = newvar()

stacklet(lgenerate, 0, Y)
stacklet(lsum, Y, 0, 10, T)
assert T == 45
```

The *future* statement immediately returns a restricted logic variable, which is bound only when the thunk it got as first argument has returned, yielding an actual return value. It is restricted in the sense that the thread which asks for a future can not assign it. [Futures](#) are implemented as a specialization of logic variables.

Note that in the current state of PyPy, we deal with coroutines, not threads (thus we can't rely on preemptive scheduling to lessen the problem with the eager consumer/producer program).

Finally, we observe that the bind, wait pair of operations is quite similar to the asynchronous send, blocking receive primitives commonly used in message-passing concurrency (like in [Erlang](#) or just plain [Stackless](#)).

## 4.3 Table of Operators

Logic variables support the following operators:

Variable Creation:

**newvar/0** nothing -> logic variable

Predicates:

**is\_free/1** any -> bool

**is\_bound/1** any -> bool

**alias\_of/2** logic variables -> bool

Mutators:

**bind/2** logic variable, any -> None | RebindingError | FutureBindingError

**unify/2** any, any -> None | UnificationError



Dataflow synchronization (blocking operations):

```
wait/1 value -> value | AllBlockedError
wait_needed/1 logic variable -> logic variable
```

## 4.4 Implementation Aspects

As above, the distinctive features of WP09 are grouped in a special object space, the Logic object space. This was motivated by the implementation of Logic variables. All other functionality is provided as standard PyPy builtins.

The most glaring feature, syntactically-wise, of the logic object space, is the half transparency of the logic variables. Basically all operations on application-level values are wrapped in calls to the *wait* generic function (or set of multi-methods), which then appropriately dispatches on the right action. One could reasonably argue that the run-time price is too high, and that it is not *pythonic* (the fourth, lesser known, theological virtue) to make such a feature implicit in Python (the classical Python mantra being *explicit is better than implicit*).

A specific type *W\_Var* is introduced at the interpreter level. This type is never accessible as such from the application level. It is only provided as an anchor for the new builtins. Thus, *bind* and *unify*, for instance, dispatch on their arguments being normal Python values or logic variables, and can hence provide adequate behaviour.

We used the interpreter-level implementation of multi-methods already used in other parts of PyPy to implement many of the builtins. Having multi-methods as an interpreter-level implementation device proved to be very convenient.

## 5 Programming With Search

### 5.1 Non-Deterministic Logic Programs

#### 5.1.1 Choice points

A logic program in the extended PyPy will be any Python program making use of the 'choice' operator. The program's entry point must be a zero arity procedure, and must be given to a solver.

The Python grammar needs only be extended like this to support the choice operator:

```
choice_stmt: 'choice:' suite ('or:' suite)+
```

For instance:

```
def foo():
    choice:
        return bar()
    or:
        from math import sqrt
        return sqrt(5)

def bar():
    choice: return -1 or: return 2
```



```
def entry_point():
    a = foo()
    if a < 2:
        fail()
    return a
```

When we encounter a choice, one of the choice points ought to be chosen non-deterministically. That means that the decision to choose does not belong to the local program (we call this *don't care non-determinism*) but to another program interested in the outcomes, typically a solver exploring the space of the logic program outcomes. The solver can use a variety of search strategies, for instance depth-first, breadth-first, discrepancy search, best-search, A\* ... It can provide just one solution or some or all of them. It can be completely automatic or end-user driven (like the [Oz explorer](#)).

Thus the program and the methods to extract information from it are truly independent, contrarily to Prolog programs which are hard-wired for depth-first exploration.

The above program contains two choice points. If given to a depth first search solver (this is the recommended option if one wants to execute a program with Prolog-like semantics, the more fancy solvers are interesting in the case of constraint solving), it will produce three spaces; the first space will be failed and thus discarded, the second and third will be solution spaces. An illustration:

```
entry_point -> foo : choice
                  /\
                 /  \
                /    \
               /      \
              /        \
             /          \
            /            \
           /              \
          /                \
         /                  \
        /                    \
       /                      \
      /                        \
     /                          \
    /                            \
   /                              \
  /                                \
 /                                  \
/                                  \
-1                                sqrt(5) (solution)
(failure) (solution)
```

To allow this de-coupling between program execution and search strategy, the computation space concept comes handy. Basically, choices are exposed in speculative computations which are embedded, or encapsulated in the so-called computation space. An intuitive approximation of computation spaces is the idea of forking a program on choices: each forked process runs in an independent address space. Solutions get out as fresh values.

Let us introduce the space method behind choice:

```
choose/1
```

A call to choose blocks the space until a call to another method, commit, is made from the solver. The choice operator can be written straightforwardly in terms of choose:

```
def bar():
    choice = choose(3)
    if choice == 1:
        return -1
    elif choice == 2:
        return 2
```

Choose is more general than choice since the number of branches can be determined at runtime. Conversely, choice is a special case of choose where the number of branches is statically determined. It is thus possible to provide syntactic sugar for it.



### 5.1.2 Driving logic programs with computation spaces

The search engine drives the computation by selecting which branches to run when the space waits on *choose* calls. To do so, the solver must peruse two computation space methods, *ask* and *commit*. Thus *ask*/0, *choose*/1 and *commit*/1 form a three-way protocol that works exactly as follows. *Ask* waits for the space to be *stable*, that is when the program running inside is blocked on a *choose* call and no other thread runs in the space. It returns either:

- 0, meaning that the space is *failed* (a space can fail due to a constraint propagator emptying one constraint variable's domain, an exception bubbling up to its entry point, or an explicit call to *fail*),
- 1, meaning that the space is *entailed* (it contains a solution to the constraint or logic program running inside),
- the value provided to *choose* by the encapsulated program. The search engine can then, depending on its strategy, *commit* the space to one choice, thus allowing the chooser to return a value which represents the choice being made.

The search and the embedded computation run in different threads. We need at least one thread for the search engine and one per space. Finally, one can *clone* computation spaces at choice points between *choose*/*commit* pairs, in the searching thread. It allows to save the current computation state before it is committed, thus avoiding the costs of full recomputations whenever *commit* sequences happen to yield failed spaces.

A small example showing a relational append:

```
def append(A, B, C):
    choice:
        unify(A, None)
        unify(B, C)
    or:
        As, Bs, X = newvar(), newvar(), newvar()
        unify(A, (X, As))
        unify(C, (X, Cs))
        append(As, B, Cs)
```

Solutions can be found as below:

```
X, Y = newvar(), newvar()
def in_space():
    return append(X, Y, (1, (2, (3, None))))

s = newspace(in_space)

from solver import solve
for sol in solve(lambda : append(X, Y, [1, 2, 3])):
    assert sol in ((None, [1, 2, 3]),
                   ([1], [2, 3]),
                   ([1, 2], [3]),
                   ([1, 2, 3], None))
```



## 5.2 Constraint programs

The logic object space comes with a modular, extensible constraint solving engine. While regular search strategies such as depth-first or breadth-first search are provided, you can write better, specialized strategies (an example would be best-first search). This section describes how to use the solver to specify and get the solutions of a constraint satisfaction problem, and then highlights how to write a new solver.

Let us mention some possible kinds of solvers:

- basic : takes no argument, enumerates all solutions;
- lazy search (this one is actually the default one since in Python writing a generator is trivial);
- general purpose (variable recomputation distance -- to tune tradoff of time and memory use), asynchronous kill to stop infinite search;
- parallel search: will spread the search on a set of machines, yielding linear speedups;
- explorer search, using a concurrent GUI engine to help a user drive the search;
- orthogonally to all the previous, it is of course possible to devise any suitable strategy (depth-first search, best-search, A\* ...).

Note that the most significant features of the constraint engine are also available in the standard object space: the [Logilab constraint solver](#) has been partly ported to RPython and can be used in standard (as in without *stackless*, the *gc framework* and constraint/logic variables) PyPy builds. This package is described in its own documentation. We will describe here the version provided in the logic object space. Both versions share the most performance-critical bits written as a pure RPython library, with respect to constraint propagation.

## 5.3 Using the constraint engine

### 5.3.1 Specification of a problem

A constraint satisfaction problem (CSP) is defined by a triple (X, D, C) where X is a set of finite domain variables, D the set of domains associated with the variables in X, and C the set of constraints, or relations, that bind together the variables of X.

So we basically need a way to declare variables, their domains and relations; and something to hold these together. The later is what we call a “computation space”. The notion of computation space is broad enough to encompass constraint and logic programming, but we use it there only as a box that holds the elements of our constraint satisfaction problem.

A problem is a zero-argument procedure defined as follows:

```
def simple_problem():
    x = domain(['spam', 'egg', 'ham'], 'x')
    y = domain([3, 4, 5], 'y')
    tell(make_expression([x,y], 'len(x) == y'))
    return x, y
```

We must be careful to return the set of variables whose candidate values we are interested in. Note that the `domain/1` primitive returns a constraint variable (which is a logic variable constrained by a domain).





## 5.3.2 Getting solutions

Now to get and print solutions out of this, we can:

```
from constraint.solver import solve
space = newspace(simple_problem)

for sol in solver.solve(space):
    print sol
```

The builtin solve function is a generator, producing the solutions as soon as they are requested.

## 5.4 Table of Operators

Note that below, “variable/expression designators” really are strings.

Space creation:

**newspace/0** nothing -> space

Finite domain creation:

**domain/2** list of any, var. designator -> Constraint variable

Expressions:

**make\_expression/2** list of var. designators, expression designator -> Expression

**AllDistinct/1** list of var. designators -> Expression

Space operations:

**newspace/1** zero-arity function -> None

**tell/1** Expression -> None

**ask/0** nothing -> a positive integer i

**choose/1** int -> None | AssertionError ('space is finished')

**clone/1** space -> space

**commit/1** integer in [1, i] -> None

**merge/1** space -> list of values (solution)

## 5.5 Extending the Search Engine

Most of this part was inspired by Christian Schulte PhD Thesis, *Programming Constraint Services* [PCS00]. The first part of the thesis, up to chapter 9, has been a driving guide for our work. The later chapters, where computation spaces are made into composable constraint combinators, have been left out. There are two reasons for this: the implementation is quite involved, and it has been shown to perform not as well as other techniques (especially with respect to techniques yielding constraint handling rules capabilities).

Hence, the semantic of the *merge* computation space operator is merely a way to extract solutions, whereas it could have been, following Schulte's steps, a merging of two spaces's state (for instance, the binding in a space of unbound variables captured and assigned to into a subspace).



### 5.5.1 Writing a solver

Here we show how the additional builtin primitives allow you to write, in pure Python, a very basic solver that will search depth-first and return the first found solution.

As we have seen, a CSP is encapsulated into a first class *computation space*. Let us see some code driving a binary depth-first search:

```
1  def first_solution_dfs(space):
2      status = space.ask()
3      if status == 0:
4          return None
5      elif status == 1:
6          return space.merge()
7      else:
8          new_space = space.clone()
9          space.commit(1)
10         outcome = first_solution_dfs(space)
11         if outcome is None:
12             new_space.commit(2)
13             outcome = first_solution_dfs(new_space)
14         return outcome
```

This recursive solver takes a space as its argument, and returns the first solution or None. In such code, the space will remain explicit as it is a first class entity manipulated from 'the outside'. Let us examine it piece by piece and discover the basics of the solver protocol.

The first thing to do is *asking* the space about its status. This may force the space to check that the values of the domains are compatible with the constraints. Every inconsistent value is removed from the variable domains. This phase is called *constraint propagation*. It is crucial because it prunes as much as possible of the search space. Then, the call to ask returns a non-negative integer value which we call the space status; at this point, all (possibly concurrent) computations happening inside the space are terminated.

Depending on the status value, either:

- the space is failed (status == 0), which means that there is no combination of values of the finite domains that can satisfy the constraints,
- one solution has been found (status == 1): there is exactly one valuation of the variables that satisfy the constraints,
- several branches of the search space can be taken (status represents the exact number of available alternatives, or branches).

Now, we have written this toy solver as if there could be a maximum of two alternatives. This assumption holds for the `simple_problem` we defined above, where a binary *distributor* (see below for an explanation of this term) has been chosen, but not in the general case. For constraint programming, it turns out that n-ary search has better performance characteristics when n is 2 [PCS00]. However, any logic program running in a space could have choice points of varying arity. A robust solver should accept that.

In line 8, we take a clone of the space; nothing is shared between space and newspace (the clone). We now have two identical versions of the space that we got as parameter. This will allow us to explore the two alternatives. This step is done, line 9 and 13, with the call to commit, each time with a different integer value representing the branch to be taken. The rest should be sufficiently self-describing.



## 5.5.2 Using distributors

A computation space might contain one thread making calls to *choose*, and there can be only one such thread per space (not respecting this practice could lead quickly to insane, undebuggable programs). We call this thread the *distributor*.

In the case of a CSP, the distributor is a simple piece of code, which works only after the propagation phase has reached a fixpoint. Its distribution policy will determine the fanout, or branching factor, of the current computation space (or node in the abstract search space).

Here are two (shipped) examples of distribution strategies:

- *AllOrNothing*, which distributes domains by splitting the smallest domain in two new domains; the first new domain has a size of one, and the second has all the other values,
- *DichotomyDistributor*, which distributes domains by splitting the smallest domain in two equal parts (or as equal as possible).

There are a great many ways to distribute... Some of them perform better, depending on the characteristics of the problem to be solved. But there is no single preferred distribution strategy. Note that the first strategy given as example there is what is used (and hard-wired) in the [MAC97] algorithm.

To use one of the built-in distributors for a constraint problem, you need to state, in the procedure that defines the problem:

```
distribute('dichotomy')
```

It is also possible to write a custom, pure-python distributor, albeit some application-level support for this might be currently missing (but could be easily added if the need arose).

## 5.6 Implementation Aspects

The non-deterministic part of this work peruses the cloning facility of PyPy's garbage collection infrastructure [D07.4]. It was shown by the Oz group that cloning was not much less efficient than trailing (as used classically in Logic languages implementations); it is even competitive with trailing [TVC99] in the case of constraint solving if cloning is mixed with recomputation steps, so as to balance the CPU/Memory consumption trade-off. Of course, cloning inherently opens the ability to distribute the load amongst several CPUs (yielding in principle linear performance increases) whereas trailing does not support parallelism at all.

## 5.7 History of the Implementation, Credits

This work was done by Strakt, Merlinux, DFKI and Logilab.

Anders Lehmann and Aurélien Campéas wrote a first pure-Python emulation of logic variables, limited computation space for the constraint solver, so as to discover the shape of the problems.

At the end of the meeting with the Oz team in Belgium, Samuele Pedroni and Carl Friedrich Bolz wrote the logic variable proxying code exploiting PyPy object spaces flexibility, along with some threading capabilities to illustrate the dataflow behaviour.

This work has been completed with all deterministic Logic operators, a specific interpreter-level scheduler, and support of non-deterministic computation (with computation space), by Aurélien with the assistance of Anders Lehman, Alexandre Fayolle and Ludovic Aubry. The initial pure-python constraint solver has been partly ported



to RPython, partly rewritten (with the help of Anders, Alexandre and Ludovic). Key low-level cloning functionality in PyPy's garbage collector framework was provided by Armin Rigo and others.

Unfortunately, as of this report, the key functionality of space cloning still doesn't work. It is quite daunting to debug it, because the full debugging cycle involves making changes to the source, compiling a whole pypy with the logic object space enabled (almost two hours on a fast machine), and then debugging the result, i.e. stepping the generated C sources (several millions lines of code, even without a backend optimization like inlining) that are, due to the implementation of the GC framework, riddled with operations that push and pop variables on a stack of reachable objects.

## 6 Semantic Web: Reasoning in OWL

OWL (the Web ontology Language) is a knowledge representation language with well-defined formal properties. It is designed for applications that need to process meaning rather than media. This way, OWL can form one of the corner stones of the Semantic Web, which is expected to allow application-independent, simultaneous access to a multitude of data, and to relationships between them.

The [Semantic Web](#) is a W3C initiative to provide means to make standards for providing data descriptions including machine readable semantics. This would allow computers to reason about the data. Quite differently the current WWW almost exclusively deals with the structure of data.

Section 1.2 of [\[OWL04\]](#) summarises the standards and tools envisaged for the Semantic Web and describes their function and interrelationships:

- [XML](#) provides a surface syntax for structured documents, but imposes no semantic constraints on the meaning of these documents.
- [XML Schema](#) is a language for restricting the structure of XML documents.
- [RDF](#) is a simple data model for referring to objects ("resources") and how they are related. An RDF-based model can be represented in XML syntax.
- [RDF Schema](#) is a vocabulary for describing properties and classes of RDF resources, with a semantics for generalization-hierarchies of such properties and classes.
- OWL adds more vocabulary for describing properties and classes: among others, relations between classes (e.g. disjointness), cardinality (e.g. "exactly one"), equality, richer typing of properties, characteristics of properties (e.g. symmetry), and enumerated classes.

OWL has three increasingly-expressive sublanguages: OWL-Lite, OWL-DL, and OWL-Full. In the present context we deal with OWL-DL, as OWL-Full is undecidable.

### 6.1 Designing an OWL-DL Reasoner

In PyPy, an OWL-DL reasoner was developed using the constraint solving package from Logilab and the Python RDF parser from [rdflib](#).

This approach is different from other approaches such as [Pellet](#) [\[POR06\]](#) and [RacerPro](#), which use the [tableaux algorithm](#).

Our idea is to convert the OWL ontology into a constraint problem, solvable by a constraint solver. To that end the OWL ontology is parsed into RDF triples that are then transformed into variables, domains and constraints utilising the semantics of OWL. The variables of the constraint problem are then the OWL classes, OWL properties and OWL individuals.



For each of the variables a domain of possible values is given. If the ontology does not explicitly supply values for a variable (i.e. by using a `owl:oneOf` predicate) the collection of all Individuals (the extension of `owl:Thing`) is chosen. This collection is built during the parsing of the RDF triples. For each triple the semantics of the predicate is used to add constraints on the variables (the “subject” and “object”). When all the triples in the ontology have been processed, the ontology can be checked for consistency by invoking the constraints. The constraints are used to narrow the domains down to the smallest number of items that satisfy the constraints. If all the domains contain at least one element, the ontology is consistent. If, on the other hand, a domain becomes empty, or a fixed size domain (created by a `oneOf` predicate) is being reduced, the ontology is not consistent, and the constraint application leads to a failure.

After a consistent ontology has been converted into a constraint problem, search can be performed by adding new constraints representing a search query.

## 6.2 Implementation of Builtin OWL Types

### 6.2.1 `owl:Class`

A `owl:Class` is defined by the members of the class. It is implemented as a class (called `ClassDomain`), which is a subclass of `AbstractDomain`.

### 6.2.2 `owl:Thing`

All individuals in OWL are a subclass of `owl:Thing`. So the domain of `owl:Thing` contains all individuals. Technically it is implemented as a subclass of `ClassDomain`.

### 6.2.3 `owl:Nothing`

`owl:Nothing` is the empty set.

### 6.2.4 `rdf:Property`

Properties are relationships between Things (or Classes). `rdf:Property` is implemented as a subclass of `AbstractDomain`, which records the subject and the object in a tuple. Internally the domain is implemented as a dictionary of a subject to a list of objects. The `getValues` method will deliver a list of tuples (or rather a generator, which produces the tuples) of subject, object.

There is also a method called `getValuesPrKey`, which produces the list of objects for a given subject.

### 6.2.5 `owl:FunctionalProperty`

`FunctionalProperty` is implemented as a subclass of `Property`, with a `FunctionalCardinalityConstraint`. This enforces the check for only one value (object) for each subject of the property.

### 6.2.6 `owl:InverseFunctionalProperty`

`InverseFunctionalProperty` is implemented as a subclass of `Property`, with an `InverseFunctionalCardinalityConstraint`. This enforces the check for only one subject for each object of the property.



### 6.2.7 owl:TransitiveProperty

TransitiveProperty is implemented as a subclass of Property. When a subject and object relationship is added to a transitive the list of pairs in the property is augmented by the transitive closure of the new pair. That means, that if the property contains (a, b) and (b, c), (a, c) is added too the list of pairs.

## 6.3 Implementation of Builtin OWL Predicates

### 6.3.1 owl:oneOf

There are two versions of owl:oneOf. The first defines a class by enumerating all the members of the class. The class is completely defined by this enumeration.

The second version defines a datatype by enumerating the values of the datatype.

The implementation creates a domain for the class, parses the elements of the list (which is the object of the oneOf triple), and fills the elements into the domain. The class is marked as final.

### 6.3.2 rdf:type

As OWL is an extension of RDF, the reasoner has to be able to support RDF semantics. For rdf:type one has to distinguish between two cases. If the object of is not a builtin OWL type, the meaning of rdf:type is implemented as a MemberConstraint that will raise an exception if the subject is not in the domain of the object. The MemberConstraint is special because it does not constrain the domain.

### 6.3.3 owl:equivalentProperty

The equivalentProperty is implemented by a special constraint that ensures that the domain of both properties are identical. This constraint does not narrow the domains of either property.

### 6.3.4 owl:subClassOf

The owl:subClassOf is implemented as a constraint for the subject class, that will remove all individuals that is not part of the object class. This means that the object class has to be defined, either by a oneOf statement or by having the defining constraints run for the object. To ensure that other defining constraints arerun first the cost attribute of a subClassOf constraint has a highvalue.

### 6.3.5 owl:equivalentClass

The owl:equivalentClass states that the members of two classes shall be the same. This is implemented with two subClassOf statements.

## 6.4 Example of Using the OWL Reasoner

Here is an example of using the reasoner on an existing ontology:

```
O = Ontology() # Create a new ontology
O.load_file("ontology.rdf") # parse the ontology into triples
O.attach_fd() # Attach finitedomains and create constraints
O.consist[LT World]_[LT World]_ency() # Check consistency
```



If the ontology is not consistent an exception is raised.

To search the ontology more constraints are added. For example to search for all individuals that have a specific value ('Green') for a property ('color'), you would do the following:

```
result = URIRef('result')
O.type(result, Restriction)
O.onProperty(result, 'color')
O.hasValue(result, 'Green')
O.consistency()
for res in O.variables[result]:
    print res
```

## 7 Using the Results for Semantic Web Applications

### 7.1 Querying the Language Technology World

A promising opportunity of applying the above results to real-world problems opened up when the developers of LT World, the world's largest information portal on Speech and Language Technologies, expressed interest in adding to the portal more generic search functionality. LT World has been developed by, and is maintained at, DFKI (<http://www.lt-world.org>). The platform informs about scientists, companies, institutions, projects, products, patents etc. and news related to speech and language technologies. The conceptual basis is entirely realized in OWL.

The eventual goal is to render LT World more transparent by retrieving snapshots of the content in a natural way, namely by querying the system using typed human language. For instance, a question like "Which institutions cooperated in Verbmobil?" will have as an answer a list of partners of the project named Verbmobil. The intended users range from knowledge engineers to end users of the portal. One requirement therefore is an appropriate time until the answer is available. Clearly this is a challenge for large ontologies; LT World has about 12 MB of RDF triples.

The problem can be divided into three subproblems:

1. Question analysis: Interpreting the human language question in terms of the ontology
2. Query answering: Generating the answer in terms of the ontology
3. Answer presentation: Presenting the answer to the user in natural language, or other media (tables)

Question analysis and answer presentation are merely linguistic problems that are outside the scope of this project. Solutions applying a grammar-based approach can be found in, e.g., [QSK05]. The result of question analysis is a semantic representation of the question that needs to be related to the vocabulary of the ontology. In a question like "What relations are there between Hans Uszkoreit and Semantic Web technologies?", it is a priori unclear what type of object "Hans Uszkoreit" and "Semantic Web technologies" are, and hence how the query should be formulated. This is solved by establishing a mapping between each name in the ontology and its natural language representations. "Semantic Web technologies" is an expression verbalizing `ltw:Semantic_Web`. Tools supporting this mapping are available (cf. [ONE06]).

The second subproblem, query answering, can be addressed by the constraint resolution facilities developed in PyPy. It is described in more detail below.

A traditional approach to query answering is to interpret the query in such a way that it guides navigation through the ontology and identifying objects related to the answer. It must be known in advance which piece of information resides in which part of the ontology, comparably to accessing a relational database. Since an application



like question answering should be designed in such a way that it can deal with different ontologies, it is desirable to keep the interface between question analysis and query answering as simple as possible. Using constraint resolution for query answering seems to be promising, as only the vocabulary representing the objects and relations in the ontology must be known, whereas “navigation in the ontology” is not required any more.

In an attempt to implement a question answering application for LT World, the following setup was decided. The query language would be SPARQL [SPQL], which is widely used and becoming a WWW standard. Question analysis would yield SPARQL expressions representing the question in terms of the ontology. The query is transformed into constraints that are added to the repository of constraints and variables. Then this constraint problem is solved. In the course of the solving, the domains of the variables are narrowed by the constraints, thus leaving only the values which satisfy the constraints in the query. Depending on the type of the query, different strategies for answer presentation are available. For instance, in the case of “Who worked for Verbmobil?”, a set of names is appropriate, but for “How many researchers worked for Verbmobil?”, it would be more appropriate to return a number, possibly followed by the names. Counting is, however, not possible within the resolver; rather it is part of the answer presentation step.

## 7.2 Using SPARQL as a query language

Typical questions include the following (in brackets, the type of the given LT World object is given):

- What is SProUT? (system)
- Do you have any information about information extraction? (technology)
- What is NLG? (technology, abbreviated)
- I want to know which projects are working on topic detection? (technology)
- Do you have any information about Hans Uszkoreit? (person)
- How many people are working in Quetal? (project)
- What is the meaning of Machine Translation? (technology)
- Who is working in SKATE? (project)
- Do you have any information about ROSANA? (project)
- Give me a definition for relation extraction. (technology)
- Who is working on natural language generation? (technology)
- Do you have any information on Verbmobil? (project, or system)

Interpreting such questions in terms of an ontology results in an expressions of the query language. In the case of ambiguous input, such as the last of the above questions, multiple answers can be given. For instance, “The project Verbmobil ... The Verbmobil system ...” For SPARQL and the LT World ontology, a representation of “Who is conducting research on parsing?” would look as follows:

```
PREFIX ltw: <http://www.lt-world.org/ltw.owl/#>
PREFIX owl: <http://www.w3.org/2002/07/owl/#>
SELECT ?person
WHERE {
    ?person_obj owl:subClassOf ltw:Active_Person.
    ?person_obj ltw:personName ?person.
    ltw:Active_Technology ltw:hasActor ?person_obj.
```





```
?person_obj ltw:Active_Technology ?topic.  
FILTER  regex(?topic, "parsing" "i").  
}
```

A few notes on the SPARQL syntax are in order. The prefixes `ltw`, `rdf` and `owl` indicate the namespaces of LT World, RDF, and OWL respectively. A symbol prefixed by a question mark is a variable that will be substituted with an individual satisfying the relations and constraints in the “WHERE” clause.

SPARQL was chosen since it promises excellent expressive power and is used in other reasoners as well. While even [Pellet](#), which may be considered the most mature OWL-DL reasoner, does not implement the full SPARQL specification, the PyPy SPARQL implementation differs in that it offers variables at the predicate position. We may thus submit questions like “What relations are there between Hans Uszkoreit and DFKI?” On the other hand, the current implementation does not cover reasoning in the definitional knowledge (“TBox”), nor does it support all XML schema types.

Questions may differ in formulation and yet be interpreted to query the same content, e.g. “What is ...” vs. “What is the meaning of ...” vs. “What is the definition of ...”. On the other hand, the questions may of course differ with respect to the object queried. In the first case, linguistic means must be provided to account for the variety. In the second case, abbreviatory means are suitable to abstract away from the specificity. In other words, a template approach is chosen to relate sets of questions to sets of SPARQL expressions. For instance, the pattern “Who is conducting research on X?”, containing a variable `X` for any technology in LT World, may correspond to a SPARQL expression similar to the above where “parsing” is replaced by the variable.

Question analysis is thus based on template matching: this way, `X` will be instantiated by the string “parsing”, and so will the SPARQL template. Moreover, the template is associated with result retrieval methods specific for the kind of question the template can cover that convert the result into the desired output.

Note that a template is usually teared to a certain type of objects, as the conditions need to express relationships specific to the type. For instance, the above example can only be used for technology, as the LT World concept `ltw:Active_Technology` must be used. A linguistically similar question for a different type of objects, such as “Who is conducting research on EU funding?”, would have to be associated with a very different SPARQL template.

### 7.3 An interpreter for query answering

A method `query_sparql()` is defined that takes a SPARQL query as its argument, creates the constraint problem using an ontology, solves the constraint problem and returns the set of values compatible with it. As the resolution process destroys the repository, a new copy of the ontology must be made available for each query. Rather than reloading the ontology from the file and performing the consistency check each time, it is computationally cheaper to preserve a copy of the consistent repository before the resolver is applied, and to reinstall the ontology using this copy afterwards. Yet this approach is still inefficient and serves as a fallback for the time being. The following section sketches a more efficient approach.

### 7.4 Performance issues

The LT World ontology is designed with the open source ontology editor [Protege](#). It is populated through web forms, or by manual additions of the maintainer. The ontology has about 1500 classes. The OWL file is about 15 MB large.

As development was done by multiple individuals, the ontology does not fully adhere to standards. For one, many inconsistencies were detected in the process of translating the ontology into a constraint problem. Many objects remain untyped, such as numbers. These objects will be translated into strings, making them inaccessible for many type-specific operations. These problems, which are non-issues in toy domains, have not fully been solved in the LT World ontology at the time this report is issued.



For testing purposes, a consistent subset with corrected type information was used. This step proved also valuable since the consistency check of the full LT World ontology takes about 90 minutes on a standard computer. After a successful testing, more detailed assessments of the relation between the size of the ontology and the runtime are in order.

Since the eventual goal is to become more independent of the size of an ontology, tuning mechanisms must be applied. Querying a large ontology will usually only affect a small part of the objects. This can be exploited in making the copy mechanism described before more efficient. Local copies of the affected objects are made before calling the resolver, and reinstalled afterwards. Thus copying efforts are largely reduced. While this has been implemented in a first version, more tests are needed. Also measurements of the efficiency gain are still missing.

### 7.5 Installation details

The following configuration of modules should be used to test the system with Python 2.4.

- Install PyPy from <http://codespeak.net/svn/pypy/dist/>
- Install <http://codespeak.net/svn/user/arigo/hack/pypy-hack/pyontology-deps> This will load the following packages:
  - Logilab-common (<ftp://ftp.logilab.fr/pub/common>)
  - Logilab-constraint (<ftp://ftp.logilab.fr/pub/constraint/logilab-constraint-0.3.0>)
  - [rdflib](#)
  - Psycho (<http://codespeak.net/svn/psyco/dist/>)

At the same time it will avoid the installation of an outdated commercial version of VisualStudio that would otherwise be needed for compilation.

- Install pyparsing.py (available from Sourceforge.net)
- Load the ontology

### 7.6 State of implementation

The resolver has been implemented and tested on a small, manually corrected subset of the LT World ontology. Various queries produce correct answers, but we are, at the time this report is written, not satisfied with the results yet. In particular, it is currently not possible yet to assess the run time in relation to the size of the ontology under consideration. These figures are urgently needed to guide further research towards a broader usability.

To reach at such figures, we need to

- reach at an OWL version of LT World that is consistent, and devise consistent subsets of different sizes;
- broaden the set of query templates that can be evaluated; this may require debugging the resolver.

## 8 Conclusion

PyPy includes a logic and constraint programming framework, which is still work in progress. It is, however, sufficiently advanced to accommodate an OWL reasoner to be written on top of it.

Several elements proper to the PyPy architecture have made this work possible: the concept of a first class object space, which allows to wrap additional semantics around normal Python semantics (the case of logic variables),



the usage of RPython, which facilitates writing interpreter-level code (by comparison with the C language), the presence of multiple dispatch methods (that facilitates the definition of builtin operators for instance), and the aspect-oriented machinery which allowed to write garbage collector-level routines for the cloning of threads (impossible with the Böhm-Weiser GC).

We can assert without doubt that this work would have been completely impossible within CPython.

The OWL-DL reasoner in PyPy was used in a real world application in the area of query answering. Experiments with a large ontology suggest that future research should address the problems scalability. Another real world problem is the correctness of existing ontologies in terms of a standardized syntax and of consistency.

## 9 Glossary of Abbreviations

The following abbreviations may be used within this document:

### 9.1 Technical Abbreviations:

AOP	Aspect Oriented Programming
AST	Abstract Syntax Tree
CPython	The standard Python interpreter written in C. Generally known as “Python”. Available from <a href="http://www.python.org">www.python.org</a> .
codespeak	The name of the machine where the PyPy project is hosted.
CCLP	Concurrent Constraint Logic Programming.
CPS	Continuation-Passing Style.
CSP	Constraint Satisfaction Problem.
CLI	Common Language Infrastructure.
CLR	Common Language Runtime.
docutils	The Python documentation utilities.
F/OSS	Free and Open Source Software
GC	Garbage collector.
GenC backend	The backend for the PyPy translation toolsuite that generates C code.
GenLLVM backend	The backend for the PyPy translation toolsuite that generates LLVM code.
GenCLI backend	The backend for the PyPy translation toolsuite that generates CLI code.
Graphviz	Graph visualisation software from AT&T.
IL	Intermediate Language: the native assembler-level language of the CLI virtual machine.
Jython	A version of Python written in Java.
LLVM	Low Level Virtual Machine - a compiler infrastructure available from University of Illinois at Urbana-Champaign
LOC	Lines of code.
Object Space	A library providing objects and operations between them, available to the bytecode interpreter via a well-defined API.
Pygame	A Python extension library that wraps the Simple DirectMedia Layer - a cross-platform multimedia library designed to provide fast access to the graphics framebuffer and audio device.



pypy-c	The PyPy Standard Interpreter, translated to C and then compiled to a binary executable program
ReST	reStructuredText, the plaintext markup system used by docutils.
RPython	Restricted Python; a less dynamic subset of Python in which PyPy is written.
Standard Interpreter	The subsystem of PyPy which implements the Python language. It is divided in two components: the bytecode interpreter, and the standard object space.
Standard Object Space	An object space which implements creation, access and modification of regular Python application level objects.
VM	Virtual Machine.

## 9.2 Partner Acronyms:

DFKI	Deutsches Forschungszentrum für künstliche Intelligenz
HHU	Heinrich Heine Universität Düsseldorf
Strakt	AB Strakt
Logilab	Logilab
CM	Change Maker
mer	merlinux GmbH
tis	Tismerysoft GmbH
Impara	Impara GmbH

## References

- [D07.4] PyPy. *Support for massive parallelism and publish about optimisation results, practical usages and approaches for translation aspects*. PyPy EU-Report D07.4, 2006.
- [OPM95] Gert Smolka. 'The Oz programming model'. In J. van Leeuwen (ed.): *Computer science today: Recent trends and developments*. LNCS, vol. 1000, Springer, 1995.
- [ECP94] Christian Schulte, Gert Smolka and Jörg Würtz. *Encapsulated search and constraint programming in Oz*. In *Second workshop on principles and practice of constraint programming*. LNCS, vol. 874, Springer, 1994.
- [PRLG92] Alain Colmerauer and Philippe Roussel. *The birth of Prolog*. In *ACM SIGPLAN notices*, Vol. 28, 1992.
- [MAC97] Daniel Sabin and Eugene Freuder. *Understanding and improving the MAC algorithm*. In Gert Smolka (Ed.): *Principles and practice of constraint programming - CP97*, Third international conference, Linz, Austria. LNCS vol. 1330, Springer, 1997.
- [CCLP93] Frank S. de Boer and Catuscia Palamidessi. *From concurrent logic programming to concurrent constraint programming*. In G. Levi (Ed.): *Advances in Logic Programming Theory*. Oxford University Press, pages 55-113, 1993.
- [MCC59] John McCarthy. *Programs with common sense*. Proceedings of Teddington conference on the mechanization of thought processes, 1959. Available at <http://www-formal.stanford.edu/jmc/mcc59/mcc59.html> (as of May 7, 2007)



- 
- [AIJ94] Christian Bessière. *Arc-consistency and arc-consistency again*. Proceedings ECAI'94 workshop on constraint processing, Amsterdam, 1994.
  - [PVR02] Peter Van Roy, Per Brand, Denys Duchier, Seif Haridi, Martin Henz, Christian Schulte. *Logic programming in the context of multiparadigm programming: The Oz experience*. In Theory and practice of logic programming, 2002.
  - [PAR86] K. Clark and S. Gregory. *Parlog: Parallel Programming in Logic*. In ACM transactions on programming languages and systems, Vol. 8, No. 1, pages 1-49, 1986.
  - [GHC88] Kazunori Ueda. *Guarded Horn clauses: A parallel logic programming language with the concept of a guard*, In M. Nivat, and K. Fuchi (eds.): Programming of future generation computers, North-Holland, Amsterdam, pages 441-456, 1988.
  - [CPR87] Ehud Shapiro (ed.). 'Concurrent Prolog : Collected papers'. MIT Press, 1987.
  - [CCP91] Vijay A. Saraswat, Martin Rinard and Prakash Panangaden. *Semantic foundations of concurrent constraint programming*. In Conference record of the eighteenth annual ACM symposium on principles of programming languages, Orlando, Florida, pages 333-352, 1991.
  - [AKL91] Sverker Janson, Seif Haridi. *Programming paradigms of the Andorra kernel language*. In Logic programming, proceedings of the 1991 international symposium. MIT Press, San Diego, California, pages 176-186, 1991
  - [CHR98] Thom Frühwirth. *Theory and practice of constraint handling rules*. In Journal of logic programming, special issue on constraint logic programming, Vol. 37, No. 1-3, pages 95-138, 1998.
  - [TVC99] Christian Schulte. *Comparing trailing and copying for constraint programming*. In Proceedings of international conference on logic programming, pages 275-289, 1999.
  - [PCS00] Christian Schulte. *Programming constraint services*. PhD Thesis, University of the Saarland, 2000.
  - [CPM01] Tobias Müller. *Constraint propagation in Mozart*. PhD Thesis, University of the Saarland, 2001.
  - [OWL04] Deborah L. McGuinness and Frank Harmelen (eds.). *OWL web ontology language. Overview*. W3C recommendation, <http://www.w3.org/TR/2004/REC-owl-features-20040210/>, 2004.
  - [SPQL] Eric Preud'homme and Andy Seaborne (eds.). *SPARQL query language for RDF*. W3C Working Draft 4 October 2006, <http://www.w3.org/TR/rdf-sparql-query/>, 2006.
  - [POR06] Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur and Yarden Katz. *Pellet: A practical OWL-DL reasoner*. Journal of web semantics, to appear. Available at <http://www.mindswap.org/papers/PelletJWS.pdf> (as of May 7, 2007).
  - [QSK05] Anette Frank, Hans-Ulrich Krieger, Feiyu Xu, Hans Uszkoreit, Berthold Cysmann, Brigitte Jörg, and Ulrich Schäfer. *Querying structured knowledge sources*. In Proceedings of AAAI-05. Workshop on question answering in restricted domains, Pages 10--19, Pittsburgh, Pennsylvania, July, 2005.
  - [ONE06] Ulrich Schäfer. *OntoNERdIE - Mapping and linking ontologies to named entity recognition and information extraction resources*. In Proceedings of the 5th international conference on language resources and evaluation (LREC), Genova, 2006.