

Lwt user manual

Jérémie Dimino

April 19, 2010

Note: this document is a work in progress. Sections 1 and 2 are mostly ready.

Contents

1	Introduction	1
2	The Lwt core library	2
2.1	Lwt concepts	2
2.1.1	Primitives for thread creation	3
2.1.2	Primitives for thread composition	4
2.1.3	Cancelable threads	4
2.1.4	Primitives for multi-thread composition	5
2.1.5	Rules	6
2.2	The syntax extension	6
2.2.1	Correspondence table	7
2.3	Other modules of the core library	8
2.3.1	Mutexes	8
2.3.2	Lists	8
2.3.3	Data streams	8
2.3.4	Mailbox variables	9
2.3.5	Reactive programming helpers	10
3	The lwt.unix library	11
3.1	Unix primitives	11
3.2	The lwt scheduler	11
3.3	Limitations	11
3.4	The logging facility	12
3.5	Detaching computation to preemptive threads	12
4	The lwt.text library	13
4.1	Text channels	13
4.2	Terminal utilities	13
4.3	Read-line	14

5	Other libraries	14
5.1	SSL support	14
5.2	Glib integration	14

1 Introduction

When writing a program, a common developer’s task is to handle IO operations. Indeed most software interact with several different resources, such as:

- the kernel, by doing system calls
- the user, by reading the keyboard, the mouse, or any input device
- a graphical server, to build graphical user interface
- other computers, by using the network
- ...

When this list contains only one item, it is pretty easy to handle. However as this list grows it becomes harder and harder to make everything works together. Several choices have been proposed to solve this problem:

- using a main loop, and integrate all components we are interacting with into this main loop.
- using preemptive system threads

Both solution have their advantages and their drawbacks. For the first one, it may works, but it becomes very complicated to write some piece of sequential code. The typical example being with graphical user interfaces freezing and not redrawing themselves because they are waiting for some blocking part of the code to complete.

If you already wrote code using preemptive threads, you shall know that doing it right with threads is a hard job. Moreover system threads consume non negligible resources, and so you can only launch a limited number of threads at the same time. Thus this is not a real solution.

Lwt offers a new alternative. It provides very light-weight cooperative threads; “launching” a thread is a very quick operation, it does not require a new stack, a new process, or anything else. Moreover context switches are very fast. In fact, it is so easy that we will launch a thread for every system call. And composing cooperative threads will allow us to write highly asynchronous programs.

In a first part, we will explain the concepts of **Lwt**, then we will describe the many sub-libraries of **Lwt**.

2 The Lwt core library

In this section we describe the basics of `Lwt`. It is advised to start an `ocaml` toplevel and try the given code examples. To start, launch `ocaml` in a terminal or in `emacs` with the `tuareg` mode, and type:

```
# #use "topfind";;  
# #require "lwt";;
```

`Lwt` is also shipped with an improved toplevel, which supports line edition and completion. If it has been correctly installed, you should be able to start it with the following command:

```
$ ocamlfind lwt/toplevel.top
```

2.1 Lwt concepts

Let's take a classical function of the `Pervasives` module:

```
# Pervasives.input_char;  
- : in_channel -> char = <fun>
```

This function will wait for a character to come on the given input channel, then return it. The problem with this function is that it is blocking: while it is being executed, the whole program will be blocked, and other events will not be handled until it returns.

Now let's look at the `lwt` equivalent:

```
# Lwt_io.read_char;;  
- : Lwt_io.input_channel -> char Lwt.t = <fun>
```

As you can see, it does not return a character but something of type `char Lwt.t`. The type `'a Lwt.t` is the type of threads returning a value of type `'a`. Actually the `Lwt_io.read_char` will try to read a character to the given input channel and *immediately* returns a light-weight thread.

Now, let's see what we can do with a `Lwt` thread. The following code create a pipe, and launch a thread reading on the input side:

```
# let ic, oc = Lwt_io.pipe ();;  
val ic : Lwt_io.input_channel = <abstr>  
val oc : Lwt_io.output_channel = <abstr>  
# let t = Lwt_io.read_char ic;;  
val t : char Lwt.t = <abstr>
```

We should now look at the state of our newly created thread:

```
# Lwt.state t;;  
- : char Lwt.state = Sleep
```

A thread may be in one of the following states:

- **Return** *x*, which means that the thread has terminated successfully and returned the value *x*
- **Fail** *exn*, which means that the thread has terminated, but instead of returning a value, it failed with the exception *exn*
- **Sleep**, which means that the thread is currently sleeping and have not yet returned a value or an exception

The thread *t* is sleeping because there is currently nothing to read on the pipe. Let's write something:

```
# Lwt_io.write_char oc 'a';;  
- : unit Lwt.t = <abstr>  
# Lwt.state t;;  
- : char Lwt.state = Return 'a'
```

So, after we write something, the reading thread has been wakeup and has returned the value 'a'.

2.1.1 Primitives for thread creation

There are several primitives for creating *Lwt* threads. These functions are located in the module *Lwt*.

Here are the main primitives:

- **Lwt.return** : *'a* -> *'a* *Lwt.t*
creates a thread which has already terminated and returned a value
- **Lwt.fail** : *exn* -> *'a* *Lwt.t*
creates a thread which has already terminated and failed with an exception
- **Lwt.wait** : *unit* -> *'a* *Lwt.t* * *'a* *Lwt.u*
creates a sleeping thread and returns this thread plus a waker (of type *'a* *Lwt.u*) which must be used to wakeup the sleeping thread.

To wake up a sleeping thread, you can use one of the following functions:

- **Lwt.wakeup** : *'a* *Lwt.u* -> *'a* -> *unit*
wakes up the thread with a value.
- **Lwt.wakeup_exn** : *'a* *Lwt.u* -> *exn* -> *unit*
wakes up the thread with an exception.

Not that this is an error to wakeup two times the same threads. `Lwt` will raise `Invalid_argument` if you try to do so.

With these informations, try to guess the result of each of the following expression:

```
# Lwt.state (Lwt.return 42);;
# Lwt.state (Lwt.fail Exit);;
# let waiter, wakener = Lwt.wait ();;
# Lwt.state waiter;;
# Lwt.wakeup wakener 42;;
# Lwt.state waiter;;
# let waiter, wakener = Lwt.wait ();;
# Lwt.state waiter;;
# Lwt.wakeup_exn wakener Exit;;
# Lwt.state waiter;;
```

2.1.2 Primitives for thread composition

The most important operation you need to know is `bind`:

```
val bind : 'a Lwt.t -> ('a -> 'b Lwt.t) -> 'b Lwt.t
```

`bind t f` creates a thread which waits for `t` to terminates, then pass the result to `f`. If `t` is a sleeping thread, then `bind t f` will be a sleeping thread too, until `t` terminates. If `t` fails, then the resulting thread will fail with the same exception.

Similarly to `bind`, there is a function to handle the case when `t` fails:

```
val catch : (unit -> 'a Lwt.t) -> (exn -> 'a Lwt.t) -> 'a Lwt.t
```

`catch f g` will call `f ()`, then waits for its termination, and if it fails with an exception `exn`, calls `g exn` to handle it. Note that both exceptions raised with `Pervasives.raise` and `Lwt.fail` are caught by `catch`.

2.1.3 Cancelable threads

In some case, we may want to cancel a thread. For example, because it has not terminated after a timeout. This can be done with cancelable threads. To create a cancelable thread, you must use the `Lwt.task` function:

```
val task : unit -> 'a Lwt.t * 'a Lwt.u
```

It has the same semantic that `Lwt.wait` except that the sleeping thread can be canceled with `Lwt.cancel`:

```
val cancel : 'a Lwt.t -> unit
```

The threads will then fails with the exception `Lwt.Canceled`. To execute a function when the thread is canceled, you must use `Lwt.on_cancel`:

```
val on_cancel : 'a Lwt.t -> (unit -> unit) -> unit
```

Note that it is also possible to cancel a thread which has not been created with `Lwt.task`. In this case, the deepest cancelable thread connected with the given thread will be cancelled.

For example, consider the following code:

```
# let waiter, waker = Lwt.task ();;
val waiter : 'a Lwt.t = <abstr>
val waker : 'a Lwt.u = <abstr>
# let t = bind waiter (fun x -> return (x + 1));;
val t : int Lwt.t = <abstr>
```

Here, cancelling `t` will in fact cancel `waiter`. `t` will then fail with the exception `Lwt.Canceled`:

```
# Lwt.cancel t;;
- : unit = ()
# Lwt.state waiter;;
- : int Lwt.state = Fail Lwt.Canceled
# Lwt.state t;;
- : int Lwt.state = Fail Lwt.Canceled
```

By the way, it is possible to prevent a thread from being canceled, by using the function `Lwt.protected`:

```
val protected : 'a Lwt.t -> 'a Lwt.t
```

Cancelling `protected t` will have no effect.

2.1.4 Primitives for multi-thread composition

We conclude this section with multi-thread composition functions. There are three such functions in the `Lwt` module: `join`, `choose` and `select`.

The first one, `join` takes a list of threads and wait for all of them to terminate:

```
val join : unit Lwt.t list -> unit Lwt.t
```

Moreover, if one thread fails, it fails with the same exception, without waiting for the other threads to terminate.

On the contrary `choose` waits for at least one thread to terminate, then returns the same value or exception:

```
val choose : 'a Lwt.t list -> 'a Lwt.t
```

For example:

```

# let waiter1, wakener1 = Lwt.wait ();;
val waiter1 : '_a Lwt.t = <abstr>
val wakener1 : '_a Lwt.u = <abstr>
# let waiter2, wakener2 = Lwt.wait ();;
val waiter2 : '_a Lwt.t = <abstr>
val wakeneré : '_a Lwt.u = <abstr>
# let t = Lwt.choose [waiter1; waiter2];;
val t : '_a Lwt.t = <abstr>
# Lwt.state t;;
- : '_a Lwt.state = Sleep
# Lwt.wakeup wakener2 42;;
- : unit = ()
# Lwt.state t;;
- : int Lwt.state = Return 42

```

2.1.5 Rules

Lwt will always try to execute the more it can before yielding and switch to another cooperative threads. In order to make it works well, you must follow the following rules:

- do not write function that may takes time to complete without using Lwt,
- do not do IOs that may block, otherwise the whole program will hang. You must instead use asynchronous IOs operations.

2.2 The syntax extension

Lwt offers a syntax extension which increases code readability and makes coding using Lwt easier. To use it add the “lwt.syntax” package when compiling:

```
$ ocamlfind ocamlc -syntax camlp4o -package lwt.syntax -linkpkg -o foo foo.ml
```

Or in the toplevel (after loading topfind):

```

# #camlp4o;;
# #require "lwt.syntax";;

```

The following construction are added to the language:

- `lwt pattern1 = expr1 [and pattern2 = expr2 ...] in expr`
which is a parallel let-binding construction. For example in the following code:

```

lwt x = f () and y = g () in
expr

```

the thread `f ()` and `g ()` are launched in parallel and their result are then bound to `x` and `y` in the expression `expr`.

Of course you can also launch the two threads sequentially by writing your code like that:

```
lwt x = f () in
lwt y = g () in
expr
```

- `try_lwt expr [with pattern1 → expr1 ...] [finally expr']`
which is the equivalent of the standard `try-with` construction but for `Lwt`. Both exception raised by `Pervasives.raise` and `Lwt.fail` are caught.
- `for_lwt ident = exprinit (to | downto) exprfinal do expr done`
which is the equivalent of the standard `for` construction but for `Lwt`.

2.2.1 Correspondence table

You can keep in mind the following table to write code using `lwt`:

without Lwt	with Lwt
<pre>let pattern₁ = expr₁ and pattern₂ = expr₂ ... and pattern_n = expr_n in expr</pre>	<pre>lwt pattern₁ = expr₁ and pattern₂ = expr₂ ... and pattern_n = expr_n in expr</pre>
<pre>try expr with pattern₁ → expr₁ pattern₂ → expr₂ ... pattern_n → expr_n</pre>	<pre>try_lwt expr with pattern₁ → expr₁ pattern₂ → expr₂ ... pattern_n → expr_n</pre>
<pre>for ident = expr_{init} to expr_{final} do expr done</pre>	<pre>for_lwt ident = expr_{init} to expr_{final} do expr done</pre>
<pre>for ident = expr_{init} downto expr_{final} do expr done</pre>	<pre>for_lwt ident = expr_{init} downto expr_{final} do expr done</pre>

2.3 Other modules of the core library

The core library contains several modules that depend only on `Lwt`. The following naming convention is used in `Lwt`: when a function takes as argument a function returning a thread that is going to be executed sequentially, it is

suffixed with “_s”. And when it is going to be executed in parallel, it is suffixed with “_p”. For example, in the `Lwt_list` module we have:

```
val map_s : ('a -> 'b Lwt.t) -> 'a list -> 'b list Lwt.t
val map_p : ('a -> 'b Lwt.t) -> 'a list -> 'b list Lwt.t
```

2.3.1 Mutexes

`Lwt_mutex` provides mutexes for `Lwt`. Its use is almost the same as the `Mutex` module of the thread library shipped with OCaml. In general, programs using `Lwt` do not need a lot of mutexes. They are only usefull for serialising operations.

2.3.2 Lists

The `Lwt_list` module defines iteration and scanning functions over lists, similar to the ones of the `List` module, but using functions that return a thread. For example:

```
val iter_s : ('a -> unit Lwt.t) -> 'a list -> unit Lwt.t
val iter_p : ('a -> unit Lwt.t) -> 'a list -> unit Lwt.t
```

In `iter_s f l`, `iter_s` will call `f` on each elements of `l`, waiting for completion between each elements. On the contrary, in `iter_p f l`, `iter_p` will call `f` on all elements of `l`, then wait for all the threads to terminate.

2.3.3 Data streams

`Lwt` streams are used in a lot of places in `Lwt` and its sub libraries. They offer a high-level interface to manipulate data flows.

A stream is an object which returns elements sequentially and lazily. Lazily means that the source of the stream is guessed for new elements only when needed. This module contains a lot of stream transformation, iteration, and scanning functions.

The common way of creating a stream is by using `Lwt_stream.from` or by using `Lwt_stream.push_stream`:

```
val from : (unit -> 'a option Lwt.t) -> 'a Lwt_stream.t
val push_stream : unit ->
  ([ 'Data of 'a
    | 'End_of_stream
    | 'Exn of exn ] -> unit) * 'a Lwt_stream.t
```

As for streams of the standard library, `from` takes as argument a function which is used to create new elements.

`push_stream` returns a function used to push new elements into the stream and the stream which will receive them.

For example:

```

# let push, stream = Lwt_stream.push_stream ();;
val push : [ 'Data of 'a | 'End_of_stream | 'Exn of exn ] -> unit = <fun>
val stream : 'a Lwt_stream.t = <abstr>
# push ('Data 1');;
- : unit = ()
# push ('Data 2');;
- : unit = ()
# push ('Data 3');;
- : unit = ()
# Lwt.state (Lwt_stream.next stream);;
- : int Lwt.state = Return 1
# Lwt.state (Lwt_stream.next stream);;
- : int Lwt.state = Return 2
# Lwt.state (Lwt_stream.next stream);;
- : int Lwt.state = Return 3
# Lwt.state (Lwt_stream.next stream);;
- : int Lwt.state = Sleep

```

Note that streams are consumable. Once you take an element from a stream, it is removed from it. So, if you want to iterates two times over a stream, you may consider “clonning” it, with `Lwt_stream.clone`. Cloned stream will returns the same elements in the same order. Consuming one will not consume the other. For example:

```

# let s = Lwt_stream.of_list [1; 2];;
val s : int Lwt_stream.t = <abstr>
# let s' = Lwt_stream.clone s;;
val s' : int Lwt_stream.t = <abstr>
# Lwt.state (Lwt_stream.next s);;
- : int Lwt.state = Return 1
# Lwt.state (Lwt_stream.next s);;
- : int Lwt.state = Return 2
# Lwt.state (Lwt_stream.next s');;
- : int Lwt.state = Return 1
# Lwt.state (Lwt_stream.next s');;
- : int Lwt.state = Return 2

```

2.3.4 Mailbox variables

The `Lwt_mvar` module provides mailbox variables. A mailbox variable, also called a “mvar”, is a cell which may contains 0 or 1 element. If it contains no elements, we say that the mvar is empty, if it contains one, we say that it is full. Adding an element to a full mvar will block until the one is taken. Taking an element from an empty mvar will block until one is added.

Mailbox variables are commonly used to pass messages between threads.

Note that a mailbox variable can be seen as a pushable stream with a limited memory.

2.3.5 Reactive programming helpers

The `Lwt_event` and `Lwt_signal` modules provide helpers for using the `react` library with `Lwt`. Among them we have `Lwt_event.next`, which takes an event and returns a thread which will wait until the next occurrence of this event. For example:

```
# let event, push = React.E.create ();;
val event : 'a React.event = <abstr>
val push : 'a -> unit = <fun>
# let t = Lwt_event.next event;;
val t : 'a Lwt.t = <abstr>
# Lwt.state t;;
- : 'a Lwt.state = Sleep
# push 42;;
- : unit = ()
# Lwt.state t;;
- : int Lwt.state = Return 42
```

Another interesting feature is the ability to limit events (resp. signals) to occurs (resp. to changes) too often. For example, suppose you are doing a program which displays something on the screen each time a signal changes. If at some point the signal changes 1000 times per second, you probably want not to render it 1000 times per second. For that you use `Lwt_signal.limit`:

```
val limit : (unit -> unit Lwt.t) -> 'a React.signal -> 'a React.signal
```

`Lwt_signal.limit f signal` returns a signal which varies as `signal` except that two consecutive updates are separated by a call to `f`. For example if `f` returns a thread which sleep for 0.1 seconds, then there will be no more than 10 changes per second. For example:

```
let draw x =
  (* Draw the screen *)
  ...

let () =
  (* The signal we are interested in: *)
  let signal = ... in

  (* The limited signal: *)
  let signal' = Lwt_signal.limit (fun () -> Lwt_unix.sleep 0.1) signal in

  (* Redraw the screen each time the limited signal change: *)
  Lwt_signal.always_notify_p draw signal'
```

3 The `lwt.unix` library

The package `lwt.unix` contains all `unix` dependant modules of `Lwt`. Among all its features, it implements cooperative versions of functions of the standard library and `unix` library.

3.1 Unix primitives

The `Lwt.unix` provides cooperative system calls. For example, the `Lwt` counterpart of `Unix.read` is:

```
val read : file_descr -> string -> int -> int -> int Lwt.t
```

`Lwt.io` provides similar features that buffered channels of the standard library (of type `in_channel` or `out_channel`) but cooperatively.

`Lwt.gc` allow you to register finaliser that return a thread. At the end of the program, `Lwt` will wait for all the finaliser to terminates.

3.2 The `lwt` scheduler

The `Lwt.main` contains the `Lwt main loop`. It can be customized by adding filters, and/or by replacing the `select` function.

Filters are responsible to collect sources to monitor before entering the blocking `select`, then to react and wakeup threads waiting for sources to become ready.

3.3 Limitations

To avoid blocking, `Lwt` puts all file descriptors (except `stdin`, `stdout` and `stderr`) in non-blocking mode. Unfortunately no `unix` supports non-blocking mode for regular files. In practise a file descriptor for a regular file will always appears to be readable and writable even if this is not the case (for example because the file is located on a network file system and the server is not available). There is no workaround for this.

By the way, since reading/writing a file will never appears to block, reading or writing a big file may prevent other threads to run until the operation on the file are finished. To avoid this, `Lwt.io` yield every 0.05 seconds for regular files.

Since there is no support at all for asynchronously reading the contents of a directory, `Lwt.directory` uses the same trick.

3.4 The logging facility

The package `lwt.unix` contains a module `Lwt.log` providing loggers. It support logging to a file, a channel, or to the `syslog` daemon. You can also defines your own logger by providing the appropriate functions (function `Lwt.log.make`).

Several loggers can be merged into one. Sending logs on the merged loggers will send these logs to all its components.

For example to redirect all logs to `stderr` and to the syslog daemon:

```
# Lwt_log.default_logger :=
  Lwt_log.merge ~level:Lwt_log.Info [
    Lwt_log.channel
      ~level:Lwt_log.Info
      ~close_mode:'keep
      ~channel:Lwt_io.stderr ();
    Lwt_log.syslog
      ~level:Lwt_log.Info
      ~facility:'User ();
  ]
;;
```

`Lwt` also provides a syntax extension, in the package `lwt.syntax.log`. It adds the following construction to the language:

Log#level format arguments

Where *level* is one of `debug`, `info`, `warning`, `error`, `fatal`. For example:

```
Log#debug "x = %s" x
```

The advantages of using the syntax extension are the following:

- it automatically prefix the format with the module name
- it check the log level before calling the logging function, so arguments are not computed if not needed
- debugging logs can be removed at parsing time

By default, the syntax extension remove all logs with the level `debug`. To keep them pass the command line option `-lwt-debug` to `camlp4`.

3.5 Detaching computation to preemptive threads

It may happen that you want to run a function which will take time to compute or that you want to use a blocking function that cannot be used in a non-blocking way. For these situations, `Lwt` allow you to *detach* the computation to a preemptive thread.

This is done by the module `Lwt_preemptive` of the `lwt.preemptive` package which maintains a spool of system threads. The main function is:

```
val detach : ('a -> 'b) -> 'a -> 'b Lwt.t
```

`detach f x` will execute `f x` in another thread and asynchronously wait for the result.

The `lwt.extra` package provides wrappers for a few blocking functions of the standard C library like `gethostbyname` (in the module `Lwt.lib`).

4 The `lwt.text` library

The `lwt.text` library provides functions to deal with text mode (in a terminal). It is composed of three following modules:

- `Lwt_text`, which is the equivalent of `Lwt_io` but for unicode text channels
- `Lwt_term`, providing various terminal utilities, such as reading a key from the terminal
- `Lwt_read_line`, which provides functions to input text from the user with line editing support

4.1 Text channels

A text channel is basically a byte channel plus an encoding. Input (resp. output) text channels decode (resp. encode) unicode characters on the fly. By default, output text channels use transliteration, so they will not fail because text you want to print cannot be encoded in the system encoding.

For example, with your locale set to “C”, and the variable `name` set to “Jérémie”, you got:

```
# lwt () = Lwt_text.printf "My name is %s" name;;
My name is J?r?mie
```

4.2 Terminal utilities

The `Lwt_term` allow you to put the terminal in *raw mode*, meaning that input is not buffered and characters are returned as the user types them. For example, you can read a key with:

```
# lwt key = Lwt_term.read_key ();;
val key : Lwt_term.key = Lwt_term.Key_control 'j'
```

The second main feature of `Lwt_term` is the ability to print text with styles. For example, to print text in bold and blue:

```
# open Lwt_term;;
# lwt () = printlc [fg blue; bold; text "foo"];;
foo
```

If the output is not a terminal, then `printlc` will drop styles, and act as `Lwt_text.printl`.

4.3 Read-line

`Lwt_read_line` provides a full featured and fully customisable read-line implementation. You can either use the high-level and easy to use `read_*` functions, or use the advanced `Lwt_read_line.Control.read_*` functions.

For example:

```
# open Lwt_term;;
# lwt l = Lwt_read_line.read_line ~prompt:[text "foo> "] ();;
foo> Hello, world!
val l : Text.t = "Hello, world!"
```

The second class of functions is a bit more complicated to use, but allow to control a running read-line instances. For example you can temporary hide it to draw something, you can send it commands, fake input, and the prompt is a signal so it can change dynamically.

5 Other libraries

5.1 SSL support

The package `lwt.ssl` provides the module `Lwt_ssl` which allow to use SSL asynchronously

5.2 Glib integration

The `lwt.glib` provides an alternative main loop for `Lwt` which uses the glib one. This allow you to write GTK application using `Lwt`. The one thing you have to do is to call `Lwt_glib.init` at the beginning of you program.