Oct 2 and 4, 2006
Lecture 8: Threads, Contd

October 4, 2006

1 ADMIN

I have posted lectures 3-7 on the web. Please use them in conjunction with the Notes you take in class. Some of the material in these lecture notes may not be discussed in class, but you are are to understand them.

On Oct 2, we spent time going over the mechanics of doing hw2 – timing and using pipes for Communication.

READING: Start reading Chapter 6 on Process Synchronization.

2 Review

1. Q: Describe the two general methods for processes to communicate among themselves in UNIX. Discuss how to do this, and their pros/cons.

   A: The 2 methods are Pipes and Files. Pipe is a 2-way communication (there is no multi-way pipe). Moreover, the 2 communicating processes need to have a common ancestor (regard a process as its own ancestor).

   Files is a more neutral medium and can involve any number of processes which have no ancestor/descendent relation. So this is a pro for using files.

   Some synchronization is necessary to prevent simultaneous writes into the file. One standard trick is to create a file with a special name like "IwantX". If someone has already created such a file, another process can detect this, and will wait until "IwantX" does not exists. Thus we create/delete the file "IwantX" to achieve MUTEX. We can then use another file "X" to share data.

   Q: How do you establish a 2-way communication link between two child processes (called X and Y) using pipes? Go through the mechanics, but do not describe the C program. Be sure to close any I/O files that are not used.

   A: X and Y must be children of a common parent process. The parent has to set up two pipes, called XtoY and YtoX. It also has a Boolean variable called IamX. The parent will set IamX=true, and fork to create process X. Then it will set IamX=false, and fork to create process Y. Then X will close XtoY[0] and YtoX[1]. Similarly, Y will close XtoY[1] and YtoX[0]. They can now communicate with X writing into XtoY[1], and reading from YtoX[0]. Similarly for Y.

3 Timing in C Programs

You will need to use this information for hw2.

There are two units for time in C and unix: in seconds, and in clock ticks. Clock ticks has a much finer granularity than seconds. The corresponding functions to obtain such times are called `time` and `clock`. Both functions return some integer type (int, long, etc) with the following interpretations:

`time()` returns the number of seconds since Jan 1st 1970 (conventional birth moment of UNIX). ¡li¿ `clock()` returns the number of clock ticks since the execution of the program started. You may identify the clock ticks with the number of CPU cycles. To convert clock ticks into seconds, you must divide by `CLOCKS_PER_SEC`.

These functions and definitions are found in include files called time.h and sys/types.h. E.g., on my computer, the later file is in /usr/include/sys/types.h.

The function `time` takes a pointer to `time_t` as an argument and returns a value of `time_t` on exit. In simple usage, the pointer can be set to NULL, so we only use the return value.

```
/* file: mytime.c
 *
 * Operating Systems Class, Fall 2006, Professor Yap.
 *  THIS code is adapted from the website
 * http://beige.ucs.indiana.edu/B673/node104.html by Zdzislaw Meglicki
 */
#include <sys/types.h>
#include <time.h>
#include <unistd.h>
#include <stdio.h>
#include <math.h>

      main()
      {
        time_t  t0, t1;
        clock_t c0, c1;

        long count;
        double a, b, c;

        t0 = time(NULL);
        c0 = clock();

        printf ("\tbegin time in sec:          %ld\n", (long) t0);
        printf ("\tbegin clock ticks:           %d\n", (int) c0);
        printf ("\t\tsleep for 5 seconds ... \n");
        sleep(5);

        printf ("\t\tperform some computation ... \n");
        for (count = 1l; count < 10000000l; count++) {
           a = sqrt(count); b = 1.0/a; c = b - a;
        }

        t1 = time(NULL);
        c1 = clock();

        printf ("\tend time in sec:              %ld\n", (long) t1);
        printf ("\tend clock ticks:              %d\n", (int) c1);
        printf ("\telapsed elapsed time in sec:  %ld\n", (long) (t1 - t0));
        printf ("\telapsed elapsed ticks in sec: %f\n",
         (float) (c1 - c0)/CLOCKS_PER_SEC);
      }
```

Saving this program as time.c, we compile and run it on my CYGWIN:

```
yap@prog[68c] gcc time.c -lm
yap@prog[684] a.exe
using UNIX function time to measure wallclock time ...
using UNIX function clock to measure CPU time ...
```

```
        begin time in sec:        1159825726
        begin clock ticks:        61
                sleep for 5 seconds ...
                perform some computation ...
        end time in sec:          1159825738
        end clock ticks:          6374
        elapsed time in sec:      12
        elapsed ticks in sec:     6.313000
yap@prog[685]
```

REMARK: In CYGWIN, we could just call ”gcc time.c”, but in general, you need to tell the compiler to look search the math library ”-lm”. You might also try ”time a.exe” to use the UNIX time function to compare with mytime. The two results should be consistent!

4 POSIX THREADS

Threads are quite diverse, and so a generic treatment is inadequate to point out differences between the various versions.

We must go into the details of some important realizations of threads for this purpose.

We are going to look at **POSIX threads** (or “pthreads”). To ensure portability of code, you should try to program only pthreads. Note: POSIX = ”portable operation system interface” (a registered trademark of IEEE).

- Is pthreads available in cygwin?

  A: Yes! The library is found in C:/cygwin/lib/libpthread.a, and the include file in C:/cygwin/usr/include/pthread.h.

- Tutorials? A: There are various tutorials to be found on the web. One such tutorial (which we henceforth call ”pthreadTutor”) may be found at Lawrence Livermore National Lab (LLNL), at the URL
  `www.llnl.gov/computing/tutorials/workshops/workshop/pthreads/MAIN.html`

- Books? A: See ”Threads Primer: A guide to multithreaded programming” by Bil Lewis and Daniel J. Berg, Sunsoft Press, 1996.

- What are some models for writing threaded programs?

  Your program must break up its task into work that can be done by independent concurrent threads.

  (1) Manager/worker: manager thread assigns work to other worker threads. Typically, manager handles all inputs and parcels out work to workers. (can have static or dynamic worker pool) [Tanenbaum] illustrates this with web server application.

  (2) Pipeline: line an automobile assembly line.

  (3) Peer: the manager thread is like a worker thread after the initial setup.

- How can I use pthread?

  It is only defined for the C language.

  Your C program must include ”pthread.h”.

  Please make sure that your `#include <pthread.h>` statement is the very first of all your include files. REASON: other include files may have ”thread-safe” features which are turned on only after pthread.h has been included.

  When you compile and link, be sure to tell your linker to use the pthread library. E.g.,

  ```
  gcc -lpthread mythreadprog.c
  ```

- What is the pthread API?

  There are three main types of objects: threads, mutex variables, condition variables.

  But associated with each main type of objects are corresponding **attribute objects**. Thus we have objects for (resp.) thread attributes, mutex attributes and condition attributes. Some people call them **opaque objects**.

  The API has over 60 subroutines divided into 3 classes of calls:

  (1) Thread management: create, detach, join, etc. set/query thread attributes (joinable, scheduling, etc).

  (2) Mutexes: create, destroy, lock, unlocking mutexes. set/query mutex attributes.

  (3) Condition variables: communication between threads that share a mutex. Create/destroy/wait/signal. Also set/query condition variable attributes.

- Names in the pthread module can be bewildering at first, so remember the following name prefix conventions:

  ```
  pthread_... are threads themselves and misc.subroutines
  pthread_attr_ are thread attribute objects
  pthread_mutex_ are mutexes
  pthread_mutex_attr are mutex attribute objects
  pthread_cond_ are condition vars
  pthread_cond_attr are condition attribute objects
  pthread_key_ are thread-specific data keys
  ```

  In other words, every method in this API will begin with the prefix `pthread_`.

- Attribute Objects. This is an interesting concept that must be mastered if you want to use pthreads. We focus on thread attributes at first.

  A thread attribute object stores the following information (the default values are in parenthesis):
  Scope (PTHREAD_SCOPE_PROCESS)
  Detached State (PTHREAD_CREATE_JOINABLE)
  Stack Address (NULL)
  Stack Size (NULL)
  Priority (NULL)

  When we create a thread, one of the arguments is an attribute object. The attributes are thereby transferred to the thread.

- HOW TO USE ATTRIBUTE OBJECTS:

  1. You must create an attribute object before using it. Do:

     ```
     // declare attr to be attribute variable
     pthread_attr_t   attr;

     // initialize attr
     if (0 != phtread_attr_init(pthread_attr_t *attr))
          perror("initializing pthread attribute");
     ```

     This returns a 0 if successful, else an error number is returned. Now `attr` is an initial object with all the default values.

  2. You can subsequently change the attribute fields. E.g., to set the detached state, do

     `phtread_attr_setdetachstate(pthread_attr_t *attr, int state)`

     where `state=PTHREAD_CREATE_DETACHED` or `state=PTHREAD_CREATE_JOINABLE`.

3. To see the current value of the detached state, use
   `phtread_attr_getdetachstate(const pthread_attr_t *attr, int *state)`
4. If you subsequently change the attribute object, this has no effect on the thread which was created earlier with this object.
5. REMARKS ON STACK: The default is no stack – the stack address and stack size pointers are both NULL in the pthread attribute object.

   When do you need a thread-specific stack? If the thread code is run by more than one thread, you should consider giving your thread its own stack! *BE SURE THAT EACH THREAD HAS A DIFFERENT STACK ADDRESS.*

   To get the thread-specific stack you need to set the stack size and stack address attributes.
6. SETTING STACK SIZE: You might think that this is easy. But there is a catch: if you set it smaller than some constant `PTHREAD_STACK_MIN`, there will be an error (presumably the default stack size of 0 will be used). Also, I found that `PTHREAD_STACK_MIN`, was not defined in my environment. So I had to define my own value and verify that it is big enough. The following seems good:

```
#define PTHREAD_STACK_MIN 10000
size_t my_stacksize;

if (0!= pthread_attr_setstacksize(&attr, (size_t)PTHREAD_STACK_MIN))
     perror("set stack size");
pthread_attr_getstacksize(&attr, &my_stacksize);
        printf("my stack size = %d \n", (int)my_stacksize);
```

7. SETTING A STACK ADDRESS: Well, you first need to allocate the space with the standard system call ”malloc”.

```
void * stackaddr;                // stack address pointer
stackaddr = (void *)malloc((size_t)PTHREAD_STACK_MIN);

pthread_attr_setstackaddr(&attr, &my_stacksize);
```

8. SCOPE NOTE: The scope of the thread can be ”process level” (default) or ”system level”. A process-scope thread is not visible to the kernel, and a system-scope is. E.g., the system does not schedule a process-level thread.

We also say the process is **unbound** when it has has process-scope, and **bound** otherwise. This terminology arise from the fact that bound processes are identified with a LWP (**light weight process**) which are kernel objects capabable of being scheduled.

- OK, I am ready, how do I create a pthread?

  Initially, there is a single **main thread** associated with the main() program. All other threads must be explicitly created by the programmer using: Here is the exact prototype for the call:

```
int pthread_create( pthread_t *thread,
        const pthread_attr_t *attr,
        void * (*start_routine)(void *),
void *arg);
```

  Suppose we call: `pthread_create(&thread, &attr, start_routine, (void *)&arg)`.

  A successful call returns a 0. The new thread ID is pointed to by the `thread` argument. The caller can use this ID to perform other operations.

The `attr` parameter is used to set thread attributes. You can use a ”thread attribute object” or NULL to accept the defaults.

The `start_routine` is a C routine that the thread will execute when it is created. The argument to `start_routine` is a pointer to a void; this argument is the last argument of `pthread_create()`.

The last `arg` is passed to the `start_routine` by reference (recase as a pointer of type void).

- HOW TO TERMINATE THREADS?

  This can happen in several ways:

  – the main (initial) thread returns

  – the thread calls `pthread_exit`

  – it is canceled by another thread via `pthread_cancel`

  – the process is terminated by a call to exit or exec

  Note that the thread_function turns a `void *`. In particular, the routine can call `int pthread_exit(void *status)` to exit, so status contains the return value.

```
int pthread\_exit(void *status)
```

  If the main thread finishes using `pthread_exit()` before its descendent threads, the other threads continue running; if `pthread_exit()` is not used, the descendents are terminated automatically.

  The termination ”status” is stored as a void pointer which any thread may join. To understand this remark, we must know that every thread will be defined to have either the DETACHED or the JOINABLE state (but not both). If DETACHED, all its resources and exit status are discarded when the thread terminates. If JOINABLE, this information will be retained until it is ”joined” to another thread (and thereby available to this thread). The default state is JOINABLE.

```
int pthread_join (pthread_t thread, void **status_ptr);
```

  CLEANUP: `pthread_exit()` does not close files...

  SUGGESTION: Use `pthread_exit()` to exit from all threads, esp. main().

- How about a sample code?

```
#include <pthread.h>
#include <stdio.h>
#define NUM_THREADS 5

void *PrintHello(void *threadid) {
   printf("\n%d: hello world!\n", threadid);
   pthread_exit(NULL);
}

int main (int argc, char * argv[]) {
   pthread_t threads[NUM_THREADS];
   int rc, t;
   for (t=0; t<NUM_THREADS; t++) {
      printf("Creating thread %d\n", t);
      rc = pthread_create( &threads[t], NULL, PrintHello, (void *)t);
      if (rc){
         printf("ERROR; return code of pthread_create() is %d\n", rc);
 exit(-1);
      }
```

```
    }
    pthread_exit(NULL);
}
```

- Mutexes. Let us consider next the mutex synchronization objects. Each mutex is essentially an int variable that are access in a very stylized way: either **lock** or **unlock**. Let L and U be the number of lock and unlock operations on a mutex variable. If $L > U$, the mutex is in the "lock state", and exactly $L - U$ of the processes that performed the lock operation will be blocked. Any "unlock" operation will unblock one of these processes.

  ```
  int
  pthread_mutex_init ( pthread_mutex_t *mut, const pthread_mutexattr_t *attr);
      -- creates a mutex variable
  int
  pthread_mutex_lock ( pthread_mutex_t *mut);
      -- locks a mutex variable
  int
  pthread_mutex_unlock ( pthread_mutex_t *mut);
      -- unlocks a mutex variable
  int
  pthread_mutex_trylock ( pthread_mutex_t *mut);
      -- either acquires a mutex variable or returns EBUSY.
  int
  pthread_mutex_destroy ( pthread_mutex_t *mut);
  ```

- Condition Variables. Condition variables allows you to wait for some predicate to hold on the condition variables.

  It is always used in conjunction with an associated mutex variable.

  There are two main operations on a cond.variable: wait and signal.

  USAGE: this usage is very stylized (in some sense, the concept of cond.variables is not a primitive concept). Let "cond" and "mut" be the condition var. and associated mutex var.
  – HOW TO WAIT: thread locks "mut" then then waits on "cond" AND "mut". When wait returns, it should unlock "mut". [REMARK: while waiting, "mut" is actually unlocked by the wait call; it is locked again upon return from wait call.]
  – HOW TO SIGNAL: thread locks "mut" and then signals "cond". When done, it unlocks "mut".

  ```
  1. Initialisation:
  int
  pthread_cond_init (pthread_cond_t *cond, pthread_condattr_t *attr);

  -- as usual, *attr can be NULL.
  2. Waiting:
  int
  pthread_cond_wait (pthread_cond_t *cond, pthread_mutex_t *mut);
  -- caller thread always blocks.
  -- NOTE that a mutex variable (mut) is required.
  -- You MUST lock "mut" before this call, and unlock "mut" after.
  -- This call is equivalent to:
  a. pthread_mutex_unlock (mut);
  b. block_on_cond (cond);
  c. pthread_mutex_lock (mut);
  -- Note that c. requires the mutex lock to be acquired.
  ```

```
3.  Signalling:
int
pthread_cond_signal (pthread_cond_t *cond);
-- this wakes up (at least?) one thread waiting on cond.
["At least" seems wrong: I think it should be "at most".
E.g., if there are no threads waiting on cond., then
nothing is woken up, according to description of signal.
If there are many, then I think you want to wake
up only one.  OTHERWISE you should use "broadcast"
instead of "signal".]
-- because of step c. above, the threads will exit
the block, one at a time.
-- You MUST lock "mut" before this call, and unlock "mut" after.
4.  Broadcast Signalling:
int
pthread_cond_broadcast (pthread_cond_t *cond);
-- wakes up all threads blocked on cond.
5.  Waiting with timeout:
int
pthread_cond_timedwait (pthread_cond_t *cond, pthread_mutex_t *mut,
                        const struct timespec *abstime);
-- similar to wait, but with timeout (absolute time of day)
6.  Deallocation:
int
pthread_cond_destroy (pthread_cond_t *cond);
```

- Example of Use of Conditional Variables

```
        pthread_mutex_t count_lock;
        pthread_cond_t count_nonzero;
        unsigned count;

        decrement_count()
           { pthread_mutex_lock(&count_lock);

             while (count == 0)
                pthread_cond_wait(&count_nonzero, &count_lock);
             count = count - 1;
             pthread_mutex_unlock(&count_lock);
           }

        increment_count()
           { pthread_mutex_lock(&count_lock);
             if (count == 0)
               pthread_cond_signal(&count_nonzero);
             count = count + 1;
             pthread_mutex_unlock(&count_lock);
            }
```

Source: David Marshall (www.cs.cf.ac.uk/Dave/C/node31.html)

- Other useful functions:

  `pthread_t pthread_self()` returns the its own thread id.