

Aufgabe 6: palim (14.0 Punkte)

Schreiben Sie ein Programm **palim** (**Parallel Line Matcher**), welches in regulären Dateien die Anzahl der Zeilen, die eine bestimmte Zeichenkette (*string*) enthalten, ausgibt.

Das Programm wird wie folgt aufgerufen:

```
palim <string> <max-grep-threads> <trees...>
```

Die als Parameter übergebenen Verzeichnisbäume (*trees*) werden jeweils in einem eigenen Thread rekursiv durchsucht (*crawl-Thread*). Für jede gefundene reguläre Datei wird wiederum ein eigener Thread (*grep-Thread*) erzeugt (**pthread_create(3)**), welcher ihren Inhalt zeilenweise durchsucht (**fopen(3)**, **fgets(3)**, **strstr(3)**). Die Anzahl der aktiven grep-Threads wird durch einen Parameter (*maxgrep-threads*) limitiert. Falls bereits so viele Threads aktiv sind, wird vor dem Erzeugen eines weiteren Threads **passiv** gewartet, bis ein Thread terminiert.

Während der Ausführung soll das Programm folgende Informationen in einer Zeile ausgeben:

- Anzahl der Zeilen, die Treffer enthalten (*lineHits*)
- Anzahl der durchsuchten Zeilen (*lines*)
- Anzahl der Dateien mit mindestens einem Treffer (*fileHits*)
- Anzahl der durchsuchten Dateien (*files*)
- Anzahl der durchsuchten Verzeichnisse (*dirs*)
- Anzahl der aktuell aktiven grep-Threads (*grepThreads*)

Die Statusausgabe soll kontinuierlich durch den Hauptthread (*main*) aktualisiert werden sobald sich eine Information ändert. Zwischenzeitlich wartet der Hauptthread **passiv**. Hierzu ist er in geeigneter Weise mit den anderen Threads zu synchronisieren.

Die Ausgabe soll folgendes Format besitzen:

```
lineHits/lines lines, fileHits/files files, dirs directories, grepThreads active threads
```

Bei Zugriffen auf globale Datenstrukturen muss auf ausreichende Synchronisierung geachtet werden. Langsame Operationen (z.B. **printf(3)**) dürfen dabei nicht in kritischen Abschnitten ausgeführt werden.

Hinweise zur Aufgabe:

- Erforderliche Dateien: Makefile, palim.c
- Der Cursor kann durch Ausgabe des Zeichens `\r` an den Anfang der Zeile zurückgesetzt werden.
- Die Ausgabe einer Zeichenkette kann mittels **fflush(3)** erzwungen werden.
- Alle erzeugten Threads sollen im *detached-state* (**pthread_detach(3)**) ausgeführt werden.
- Zum Übersetzen des Programmes ist das zusätzliche Compiler-Flag `-pthread` notwendig.
- Eine Referenzimplementierung finden Sie in `/proj/i4sp1/pub/aufgabe6`
- Zum Testen kann palim mit den Verzeichnisbäumen `/usr/share/doc/texlive*` aufgerufen werden.
- Gehen Sie davon aus, dass die Dateien keine Zeilen enthalten, die länger sind als durch den **sysconf(3)**-Parameter `_SC_LINE_MAX` vorgegeben.
- Verwenden Sie für die Synchronisation das vorgegebene Semaphoren-Modul (`sem.o`, `sem.h`).

Hinweise zur Abgabe:

Bearbeitung: Zweiergruppen

Bearbeitungszeit: 11 Werktage (ohne Wochenenden und Feiertage)

Abgabezeit: 17:30 Uhr