Benchmarking Soar: Using Soar's Timers

Robert Wray
wrayre@acm.org

22 May 1999

MTSU
Middle Tennessee State University

Why Should I Use Soar's Timers?

Timer: stop watch

measures time elapsed (usually execution of a procedure)

- Process time (e.g., `ps`) provides only a coarse measure of cost
  - Distinguish between "reasoning" time and other functions (loading, callbacks, input, output, etc.)

    kernel time: measure of Soar reasoning time only
- Identify costly/inefficient reasoning
  - Which phase requires the most time?
  - How much time is spent processing input?
  - What is the time cost of match, chunking, etc?
- If you don't need to know the time costs of Soar's computations and your application does not demand "maximum" efficiency, you can probably ignore timers.

How Do I Use the Timers?

Two flags determine if timing data is collected and the level of detail of the statistics.

`soarkernel.h` definitions (`#define`)

- No definitions
  Soar will collect some basic timing data (version-dependent)
  Default in 7.0.x distributions
- `#define NO_TIMING_STUFF`
  Soar will not keep any timing statistics
- `#define DETAILED_TIMING_STATS`
  Soar will collect "detailed" (fine-grained) timing statistics
  Default in 8.2 distribution

Re-build Soar after defining the level of timing desired

`DETAILED_TIMING_STATS` has no effect if `NO_TIMING_STUFF` is defined.

stats with NO_TIMING_STUFF enabled

```
Soar 8.2 on newell.mtsu.edu at Thu May 13 09:48:33 1999

91 productions (79 default, 12 user, 0 chunks)
   + 0 justifications
39 decision cycles
170 elaboration cycles
833 production firings
2403 wme changes (1267 additions, 1136 removals)
WM size: 131 current, 174.488 mean, 242 maximum
```

What Will the Timers Tell Me?

- Soar6.2
  - Total execution ("**`run`**") time (not process time)
  - **`DETAILED_TIMING_STATS`**

    chunking

    match

    ownership
- Soar7.0.3/7.0.4
  - Kernel time (no callbacks, input, output)
  - Kernel time reported for individual phases
  - **`DETAILED_TIMING_STATS`** are also reported by phase
- Soar8.2
  - Soar7.0.3/7.0.4
  - **`DETAILED_TIMING_STATS`** includes a category for Soar8-specific processing

Timing Statistics: Soar6.2

```
Soar 6.2.5 on parrot.eecs.umich.edu at Sat May 15 15:31:29 1999

127 productions (100 default, 24 user, 3 chunks)
   + 0 justifications
Total cpu time: 0.169 seconds
27 decision cycles (6.244 msec/dc)
92 elaboration cycles (3.407 ec's per dc, 1.832 msec/ec)
438 production firings (4.761 pf's per ec, 0.385 msec/pf)
1100 wme changes (577 additions, 523 removals)
WM size: 54 current, 205.741 mean, 308 maximum
```

What Will the Timers Tell Me? (Soar7.0.3)

Soar7.0.3 reports **kernel time** as well as total CPU time

- Kernel time includes time for match, WME changes, chunking, ownership calculations, etc.
- Kernel time includes only time spent “run”-ning (not loading, parsing, etc.)
- Kernel time excludes:
  - Input function (e.g., tcl simulator)
  - Output function
  - Callbacks (monitors)

Kernel time is used to report frequencies

```
   Kernel CPU Time:       0.290 sec.
   Total  CPU Time:       5.870 sec.

 42 decision cycles (6.905 msec/dc)
137 elaboration cycles (3.262 ec’s per dc, 2.117 msec/ec)
```

What Will the Timers Tell Me? (Soar7.0.3)

Soar7.0.3 reports time per phase as well as total kernel time and total CPU time

- Individual timers are used for each phase

Derived totals

- Sum of the time spent in individual phases
- Example: Sum will always be ≤ `Kernel CPU Time`
  - Differences due to implementation of timers....
  - `Kernel CPU Time` used to report frequencies (eg, msec/dc)

```
                                                       Derived
Phases:      Input    Pref      W/M   Output Decision |     Totals
======================================================|===========
Kernel:      0.000   1.040    0.130    0.000    0.040 |      1.210
======================================================|===========

Kernel CPU Time:       1.220 sec.
```

What Will the Timers Tell Me? (Soar7.0.3)

`DETAILED_TIMING_STATS`

`Match`

Time for adding and deleting WMEs from the RETE

`Ownership`

Time for promoting and demoting
(determine goal to which an identifier is linked)

`Chunking`

Time for creating chunks (and justifications)

`Other`

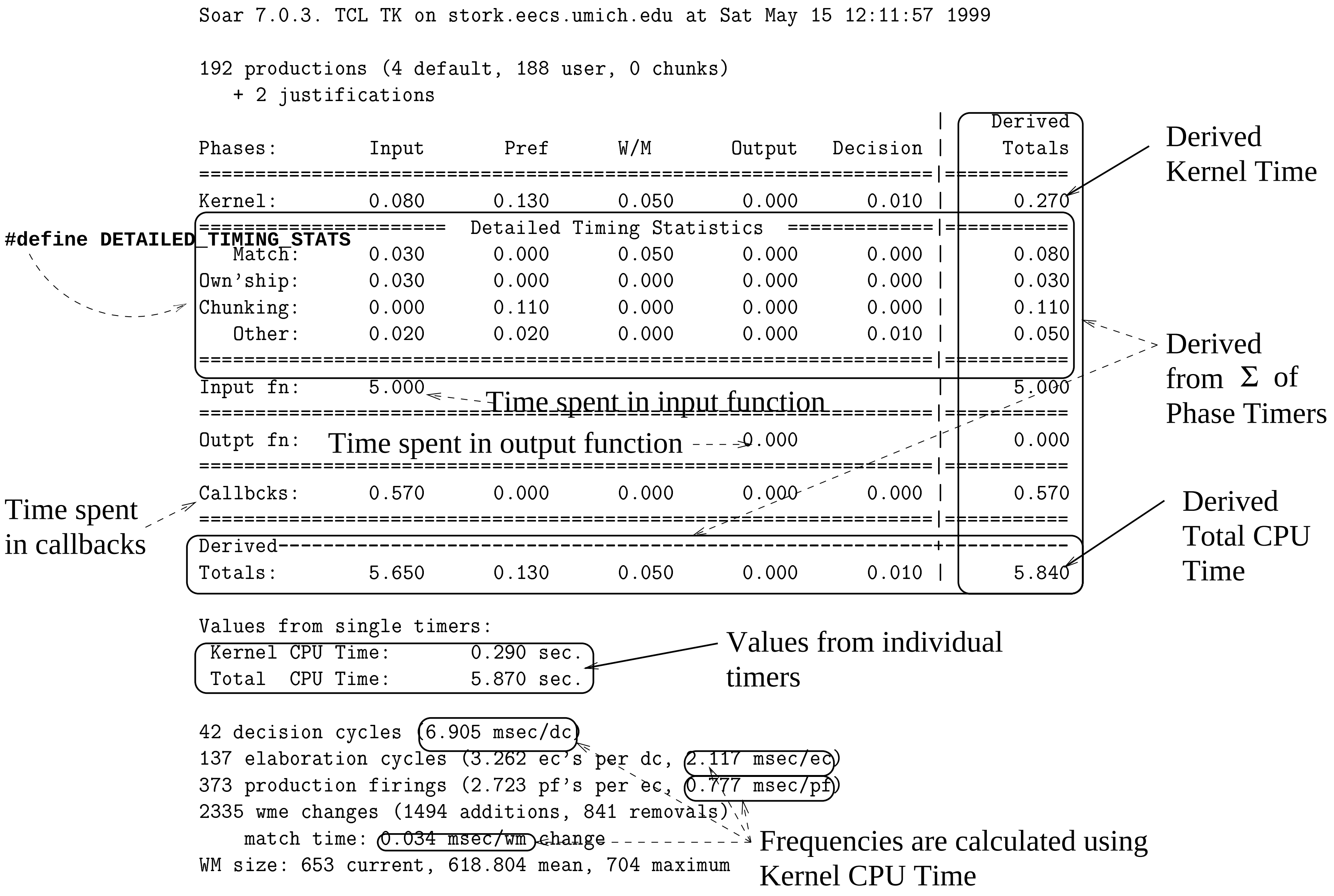
Derived total:

$$\text{Other}_{phase} = \text{KernelTime}_{phase} - (\text{Match} + \text{Ownership} + \text{Chunking})_{phase}$$

stats with DETAILED_TIMING_STATS enabled: 7.0.3

```
Soar 7.0.3. TCL TK on stork.eecs.umich.edu at Sat May 15 12:11:57 1999

192 productions (4 default, 188 user, 0 chunks)
   + 2 justifications
                                                            |    Derived
Phases:        Input       Pref      W/M       Output   Decision |     Totals
=================================================================|===========
Kernel:        0.080      0.130      0.050      0.000      0.010 |      0.270
======================  Detailed Timing Statistics  =============|===========
   Match:      0.030      0.000      0.050      0.000      0.000 |      0.080
Own'ship:      0.030      0.000      0.000      0.000      0.000 |      0.030
Chunking:      0.000      0.110      0.000      0.000      0.000 |      0.110
   Other:      0.020      0.020      0.000      0.000      0.010 |      0.050
=================================================================|===========
Input fn:      5.000                                             |      5.000
=================================================================|===========
Outpt fn:                                       0.000            |      0.000
=================================================================|===========
Callbcks:      0.570      0.000      0.000      0.000      0.000 |      0.570
=================================================================|===========
Derived----------------------------------------------------------+-----------
Totals:        5.650      0.130      0.050      0.000      0.010 |      5.840

Values from single timers:
 Kernel CPU Time:       0.290 sec.
 Total  CPU Time:       5.870 sec.

42 decision cycles (6.905 msec/dc)
137 elaboration cycles (3.262 ec's per dc, 2.117 msec/ec)
373 production firings (2.723 pf's per ec, 0.777 msec/pf)
2335 wme changes (1494 additions, 841 removals)
    match time: 0.034 msec/wm change
WM size: 653 current, 618.804 mean, 704 maximum
```

Summary: What Will the Timers Tell Me?



```
Soar 7.0.3. TCL TK on stork.eecs.umich.edu at Sat May 15 12:11:57 1999

192 productions (4 default, 188 user, 0 chunks)
   + 2 justifications
                                                                  |   Derived
Phases:         Input        Pref       W/M       Output   Decision |    Totals
==================================================================|===========
Kernel:         0.080       0.130      0.050      0.000      0.010 |     0.270
======================  Detailed Timing Statistics  =============|===========
   Match:       0.030       0.000      0.050      0.000      0.000 |     0.080
Own'ship:       0.030       0.000      0.000      0.000      0.000 |     0.030
Chunking:       0.000       0.110      0.000      0.000      0.000 |     0.110
   Other:       0.020       0.020      0.000      0.000      0.010 |     0.050
==================================================================|===========
Input fn:       5.000                                              |     5.000
==================================================================|===========
Outpt fn:                                             0.000        |     0.000
==================================================================|===========
Callbcks:       0.570       0.000      0.000      0.000      0.000 |     0.570
==================================================================|===========
Derived-----------------------------------------------------------+-----------
Totals:         5.650       0.130      0.050      0.000      0.010 |     5.840

Values from single timers:
 Kernel CPU Time:        0.290 sec.
 Total  CPU Time:        5.870 sec.

42 decision cycles (6.905 msec/dc)
137 elaboration cycles (3.262 ec's per dc, 2.117 msec/ec)
373 production firings (2.723 pf's per ec, 0.777 msec/pf)
2335 wme changes (1494 additions, 841 removals)
    match time: 0.034 msec/wm change
WM size: 653 current, 618.804 mean, 704 maximum
```

#define DETAILED_TIMING_STATS
Derived Kernel Time
Derived from Σ of Phase Timers
Time spent in input function
Time spent in output function
Time spent in callbacks
Derived Total CPU Time
Values from individual timers
Frequencies are calculated using Kernel CPU Time

What's New in Soar8.2?

New Phase: Determine Level Phase (DLP)

- Kernel phase timers include this phase
- `stats` reflects additional phase

New category in `DETAILED_TIMING_STATS`: "Operand2"

- Includes:

  Goal Dependency Set calculations

  Determination of the highest active level (DLP)

- Time measured for `Operand2` is separate from other statistics (`Operand2` time is included in `Other`)

+ Category name should be re-labeled

? Separate GDS and determine level timers?

Soar8.2 stats with NO_TIMING_STUFF disabled

```
Soar 8.2 on newell.mtsu.edu at Thu May 13 09:44:23 1999

91 productions (79 default, 12 user, 0 chunks)
   + 0 justifications
                                                                |    Derived
Phases:      Input      DLP     Pref      W/M   Output Decision |     Totals
================================================================|===========
Kernel:      0.000    0.000    1.040    0.130    0.000    0.040 |      1.210
================================================================|===========
Input fn:    0.000                                              |      0.000
================================================================|===========
Outpt fn:                                        0.000          |      0.000
================================================================|===========
Callbcks:    0.000    0.000    0.000    0.000    0.000    0.010 |      0.010
================================================================|===========
Derived---------------------------------------------------------+-----------
Totals:      0.000    0.000    1.040    0.130    0.000    0.050 |      1.220

Values from single timers:
 Kernel CPU Time:       1.220 sec.
 Total  CPU Time:       1.230 sec.

39 decision cycles (31.282 msec/dc)
170 elaboration cycles (4.359 ec's per dc, 7.176 msec/ec)
45 p-elaboration cycles (1.154 pe's per dc, 27.111 msec/pe)
833 production firings (4.900 pf's per ec, 1.465 msec/pf)
2403 wme changes (1267 additions, 1136 removals)
WM size: 131 current, 174.488 mean, 241 maximum
```

Soar8.2 stats with DETAILED_TIMING_STATS enabled

```
Soar 8.2 on newell.mtsu.edu at Thu May 13 09:52:37 1999

91 productions (79 default, 12 user, 0 chunks)
   + 0 justifications
                                                                |    Derived
Phases:      Input      DLP     Pref      W/M   Output Decision |     Totals
================================================================|===========
Kernel:      0.000    0.020    1.520    0.180    0.000    0.130 |      1.850
====================  Detailed Timing Statistics  ==============|===========
   Match:    0.000    0.000    0.000    0.150    0.000    0.000 |      0.150
Own'ship:    0.000    0.010    0.000    0.010    0.000    0.000 |      0.020
Chunking:    0.000    0.000    0.010    0.000    0.000    0.000 |      0.010
   Other:    0.000    0.010    1.510    0.020    0.000    0.130 |      1.670
Operand2:    0.000    0.010    0.000    0.000    0.000    0.000 |      0.010
================================================================|===========
Input fn:    0.000                                              |      0.000
================================================================|===========
Outpt fn:                                        0.000          |      0.000
================================================================|===========
Callbcks:    0.000    0.000    0.000    0.000    0.000    0.000 |      0.000
================================================================|===========
Derived---------------------------------------------------------+-----------
Totals:      0.000    0.020    1.520    0.180    0.000    0.130 |      1.850

Values from single timers:
 Kernel CPU Time:       1.850 sec.
 Total  CPU Time:       1.860 sec.

67 decision cycles (27.612 msec/dc)
289 elaboration cycles (4.313 ec's per dc, 6.401 msec/ec)
77 p-elaboration cycles (1.149 pe's per dc, 24.026 msec/pe)
1410 production firings (4.879 pf's per ec, 1.312 msec/pf)
4021 wme changes (2076 additions, 1945 removals)
    match time: 0.037 msec/wm change
WM size: 131 current, 174.426 mean, 241 maximum
```

How Are the Timers Implemented?

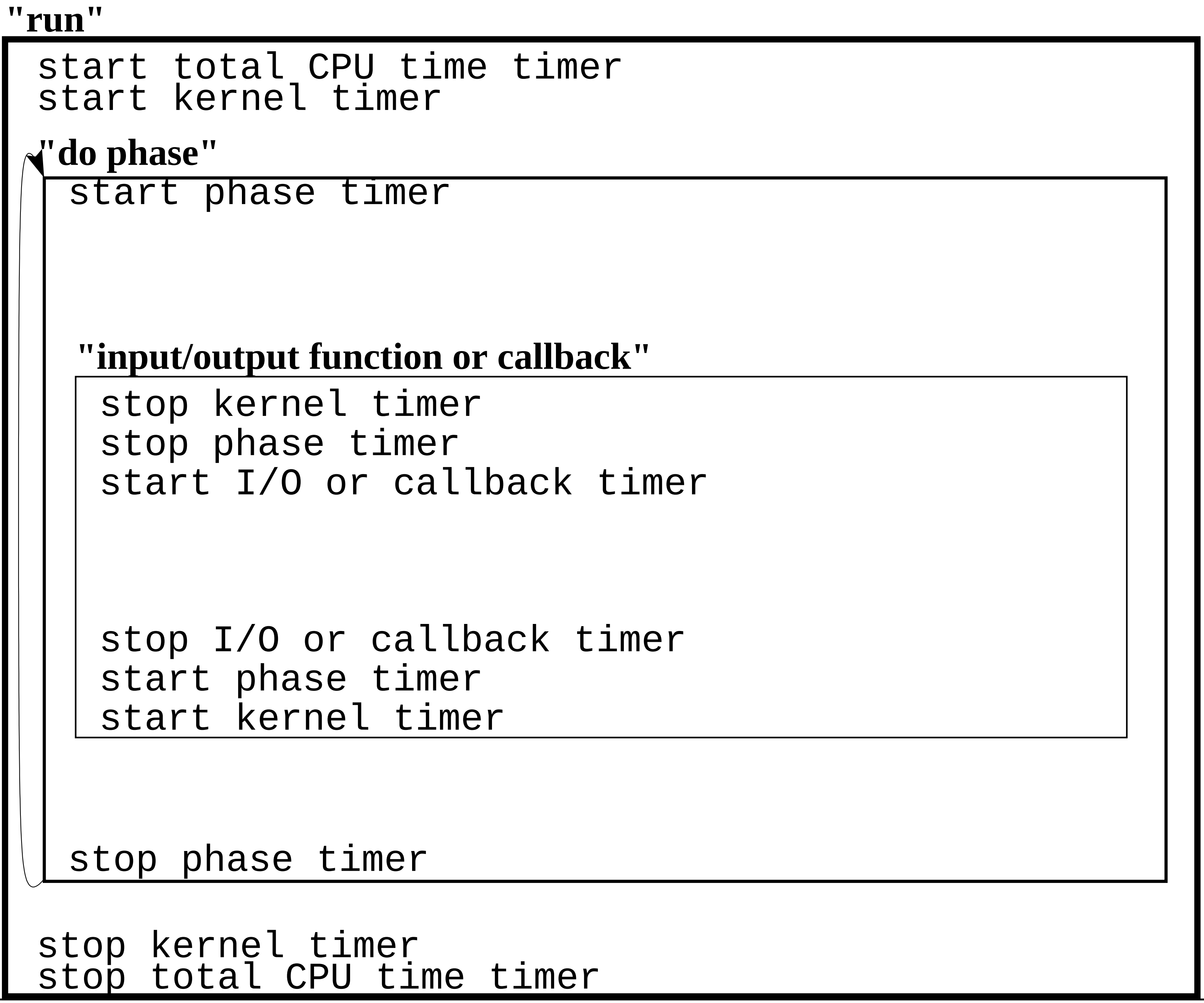
- Soar implementation provides an interface for using timers:

```
 /* Initialize/reset timer to 0 */
 void reset_timer (struct timeval *timer_to_reset);

 /* Start a timer */
 void start_timer (struct timeval *start_time);

/* Stop a timer.  Store the elapsed time since the call to start_time in
   accumulated_time */
void stop_timer(struct timeval *start_time,struct timeval *accumulated_time);

/* Return the number of seconds stored in the timer some_timer */
double timer_value (struct timeval *some_timer);
```

- Implementation of functions is dependent on operating system

    Unix: **`rusage/getrusage()`**

- Easy to add new timers using existing functions
- Users can (mostly) ignore underlying implementation
    - e.g., Users don't need to know definition of **`timeval`**



"run"

start total CPU time timer
start kernel timer

"do phase"

start phase timer

"input/output function or callback"

stop kernel timer
stop phase timer
start I/O or callback timer

stop I/O or callback timer
start phase timer
start kernel timer

stop phase timer

stop kernel timer
stop total CPU time timer

Do the Timers Have Any Limitations?

Uncertainty Principle:

Executing code to measure time of execution

- Individual calls to `start_timer` or `stop_timer` are (mostly) insignificant
- Many calls to timer functions over the course of execution

  Empirical Results (Scott Wallace):

  ∗ $\approx 1 * 10^{-4}$ seconds/timer function call

  ∗ Thousands of calls to timers for simple tasks (esp. with `DETAILED_TIMING_STATS`)

  ∗ Factor of 4 change in time data in some experiments

Recommendations:

– Use timers for development, not deployment

– Avoid using `DETAILED_TIMING_STATS` in applications with real-time demands

? Need new timing level: `KERNEL_TIME` (no phase timing)

Summary

- Timers provide information useful for measuring performance and recognizing inefficiencies
- Soar provides an interface to the timers that abstracts from the actual implementation of the timers (underlying O/S)
- Execution of timer procedures can be costly/may degrade performance

Possible Action Items

- Distribution default: No compiler directives for timing (basic timing information)
- Replace "Operand2" with a better term in Soar8.2 `stats`
- Are there other categories of calculations we should consider for the `DETAILED_TIMING_STATS`?
- Should there be another level of benchmarking? `NO_TIMING_STUFF`, **`KERNEL_TIME`**, `PHASE_TIME`, `DETAILED_TIMING_STATS`