Data for OOPSLA

Sara Achour
Martin Rinard

1 Data

1.1 Predicted Error Table

1.2 Outlier Detector Quality Table

1.3 Output Quality Table

1.4 Energy Table

1.4.1 Raw Data

1.5 Control System Based Outlier Detector

PID control:

1.5.1 Initial Step: Seeding the Outlier Detector

1. Reject the first n tasks - find the mean and standard deviation of the outputs (μ, σ)
2. update the $mean = \mu, l = \sigma, r = \sigma$. the effective bounds this detector accepts is: $[\mu - l, \mu + r]$
3. After n tasks have been used to get an initial l, r and m for the control system, we move to the control phase.

1.5.2 Control System

We have two distinct control systems, the left-bound control system, which impacts bound parameter l and the right bound control system, which impacts bound parameter r.

Continuous version:
$\mu(t) = K_p r_e + K_i \int r_e + K_d \frac{\delta}{\delta t} r_e$
Discretized version:
given new re-execution rate error r_e, previous state (r_e^*, r_d^*, r_i^*), the error, error derivative, error integral. Also given decay D=0.99.
$\mu(t) = K_p r_e + K_i(r_e + D \cdot r_i^*) + (r_e - r_e^*)$

Benchmark	Predicted Accepted Errors (p)	Actual Accepted Errors	Output	Maximum Output Error (E)	Predicted Output Error $p \cdot e$
scale	22	22	red	25	25
			green	25	25
			blue	25	25
blackscholes	22	22	price	25	25
barnes	22	22	vel[0]	25	25
			vel[1]	25	25
			vel[2]	25	25
			acc[0]	25	25
			acc[1]	25	25
			acc[2]	25	25
			siz	25	25
			phi	25	25
water/intermol	22	22	res1[0]	25	25
			res1[1]	25	25
			res1[2]	25	25
			res1[3]	25	25
			res1[4]	25	25
			res1[5]	25	25
			res1[6]	25	25
			res1[7]	25	25
			res1[8]	25	25
			res2[0]	25	25
			res2[1]	25	25
			res2[2]	25	25
			res2[3]	25	25
			res2[4]	25	25
			res2[5]	25	25
			res2[6]	25	25
			res2[7]	25	25
			res2[8]	25	25
water/poteng	22	22	res[0]	25	25
			res[1]	25	25
			res[2]	25	25

Benchmark	Accepted Correct	Rejected Correct	Accepted Error	Rejected Error	Re-Executed
scale	97.875%	0.915%	0.091%	1.118%	2.034%
blackscholes	98.611%	0.997%	0.134%	1.254%	1.254 %
water	96.578%	0.788%	1.166%	1.467 %	2.25%
barnes	96.704%	3.028%	0.134%	0.134%	3.162%

Table 1: control-based outlier detector quality

benchmark	no outlier detector	outlier detector/re-execution	outlier detector/discard
scale [PSNR]	28.1601	43.9167	23.1269
blackscholes [portfolio error]	$6.042 \cdot 10^{34}$%	$8.075 \cdot 10^{-2}$ %	0.513 %
water [position error]	nan*	$7.100 \cdot 10^{-4}$ %	$2.415 \cdot 10^{147}$%**
barnes [position error]	0.699%	0.125%	0.046%

Table 2: output quality without, with control outlier detector. (*) after the first time step, all positions go to nan. (**) spurious nans as positions, a few bad values badly skew the output.

benchmark	% main computation	% reliable re-execution	% unreliable tasks (0.85* something)	total
scale	22	23	24	25
blackscholes	32	33	34	35
water	42	43	44	45
barnes	42	43	44	45

Table 3: Energy Savings