100x100 with random base policy

avg. accumulated cost

- UCT
- AOT

avg. time per decision (milliseconds)

barto−big with random base policy

avg. accumulated cost

- UCT
- AOT

avg. time per decision (milliseconds)