Getting to Know One Another: Calibrating Intent, Capabilities and Trust for Human-Robot Collaboration (Appendix)

Joshua Lee, Jeffrey Fong, Bing Cai Kok, and Harold Soh
Dept. of Computer Science, National University of Singapore
joshua_lks@u.nus.edu, {jfong, kokbc, harold}@comp.nus.edu.sg

A. *Additional Details for TICC-MCP*

The pseudocode and a simple running example for the TICC-MCP is shown in Algorithm 1 and Fig. 1 respectively.

Algorithm 1 HA-POMCP

procedure SEARCH(h)
  repeat
    if $h = empty$ **then**
      $\bar{s} \sim \mathcal{I}$
    else
      $\bar{s} \sim B(h)$ SIMULATE($\bar{s}, h, 0$)
    end if
  until TIMEOUT
  return $\arg\max_{a^{\text{R}}} V(ha^{\text{R}})$
end procedure

procedure ROLLOUT($\bar{s}, h, depth$)
  if $\gamma^{depth} < \epsilon$ **then**
    return 0
  end if
  $a^{\text{R}}, a^{\text{H}} \sim Uniform(a^{\text{R}}), Uniform(a^{\text{H}})$
  $\tilde{a}^{\text{R}} \sim P^*(\tilde{a}^{\text{R}}|a^{\text{R}})$
  $\tilde{a}^{\text{H}} \sim P(\tilde{a}^{\text{H}}|a^{\text{H}};\psi)$
  $\bar{s}' \leftarrow T(\bar{s}, a^{\text{R}}, a^{\text{H}}, \tilde{a}^{\text{R}}, \tilde{a}^{\text{H}})$
  return $r(\bar{s}) + \gamma$ ROLLOUT($\bar{s}', ha^{\text{R}}a^{\text{H}}, depth+1$)
end procedure

procedure SIMULATE($\bar{s}, h, depth$)
  if $\gamma^{depth} < \epsilon$ **then**
    return 0
  end if
  if $h \notin T$ **then**
    return ROLLOUT($\bar{s}, h, depth$)
  end if
  $a^{\text{R}} \leftarrow \arg\max_{a^{\text{R}}} V(ha^{\text{R}}) + c\sqrt{\frac{\log N(h)}{N(ha^{\text{R}})}}$
  $a^{\text{H}} \sim \pi_H(a^{\text{H}}|V_\theta(ha^{\text{R}}a^{\text{H}}) + c\sqrt{\frac{\log N_\theta(ha^{\text{R}})}{N_\theta(ha^{\text{R}}a^{\text{H}})}})$
  $\tilde{a}^{\text{R}} \sim P^*(\tilde{a}^{\text{R}}|a^{\text{R}})$
  $\tilde{a}^{\text{H}} \sim P(\tilde{a}^{\text{H}}|a^{\text{H}};\psi)$
  $\bar{s}' \leftarrow T(\bar{s}, a^{\text{R}}, a^{\text{H}}, \tilde{a}^{\text{R}}, \tilde{a}^{\text{H}})$
  $R \leftarrow r(\bar{s}) + \gamma$ SIMULATE($\bar{s}', ha^{\text{R}}a^{\text{H}}, depth+1$)
  $B(h) \leftarrow B(h) \cup \{\bar{s}\}$
  $N(h) \leftarrow N(h) + 1$
  $N_\theta(ha^{\text{R}}) \leftarrow N_\theta(ha^{\text{R}}) + 1$
  $N_\theta(ha^{\text{R}}a^{\text{H}}) \leftarrow N_\theta(ha^{\text{R}}a^{\text{H}}) + 1$
  $V(ha^{\text{R}}) \leftarrow V(ha^{\text{R}}) + \frac{R - V(ha^{\text{R}})}{N(ha^{\text{R}})}$
  $V_\theta(ha^{\text{R}}a^{\text{H}}) \leftarrow V_\theta(ha^{\text{R}}a^{\text{H}}) + \frac{R - V_\theta(ha^{\text{R}}a^{\text{H}})}{N_\theta(ha^{\text{R}}a^{\text{H}})}$
  return R
end procedure

B. *Additional Details for Human Subject Experiments*

In human subject experiments, the shopping lists and the actual capability setup are shown in Table I. The robot running the TICC-MCP algorithm initially assumes that 1) the human has perfect capability 2) the human thinks that the robot has perfect capability.

TABLE I
SHOPPING LISTS (TOP), HUMAN CAPABILITY IN SUCCESS RATE (MIDDLE) AND ROBOT CAPABILITY IN SUCCESS RATE (BOTTOM) USED FOR EXPERIMENTS WITH HUMAN SUBJECTS.

Shopping lists

Yellow Cup	Sweet	Cleaning Agent	Green Cup
4	0	0	5
1	1	6	1
1	3	5	0
3	2	2	1
3	1	3	1
1	3	3	1
7	1	0	0
2	3	3	0

Human's capability

Yellow Cup	Sweet	Cleaning Agent	Green Cup
50%	100%	0%	100%

Robot's capability

Yellow Cup	Sweet	Cleaning Agent	Green Cup
80%	0%	100%	100%

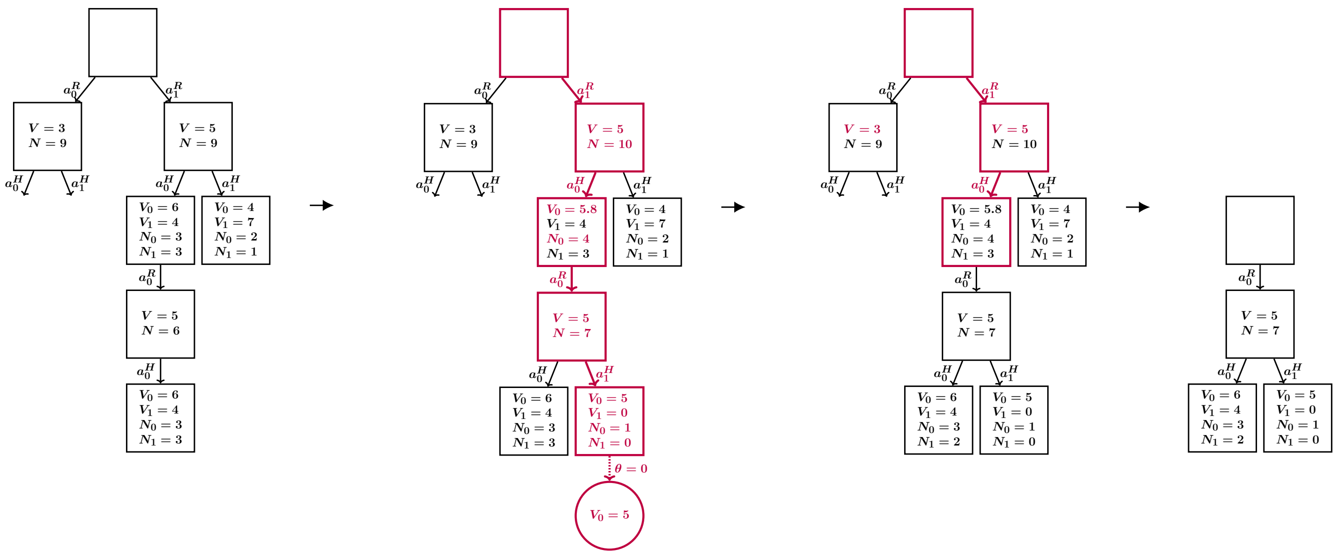

Fig. 1. An example illustration of TICC-MCP in a simple environment with 2 robot actions $a^{\mathrm{R}} \in \{a_0^{\mathrm{R}}, a_1^{\mathrm{R}}\}$, 2 human actions $a^{\mathrm{H}} \in \{a_0^{\mathrm{H}}, a_1^{\mathrm{H}}\}$, 2 reward parameters $\theta \in \{0, 1\}$, no intermediate rewards and a discount factor of 1. For brevity, the belief particles at each node are omitted. The robot constructs a search tree from multiple search iterations and stores the values V and number of visits N in every node/history (left). When visiting a new leaf node, the robot performs expansion and simulation using rollout policy, followed by backpropagation (second left). The robot compares the values in the search tree and choose action a_1^{R}; the human then chooses action a_0^{H} (second right). The robot prunes the tree and begins a new search from the updated history $ha_1^{\mathrm{R}}a_0^{\mathrm{H}}$ (right).