A Game-Theoretic Approach for Adaptive Action Selection in Close Proximity Human-Robot-Collaboration (HRC)

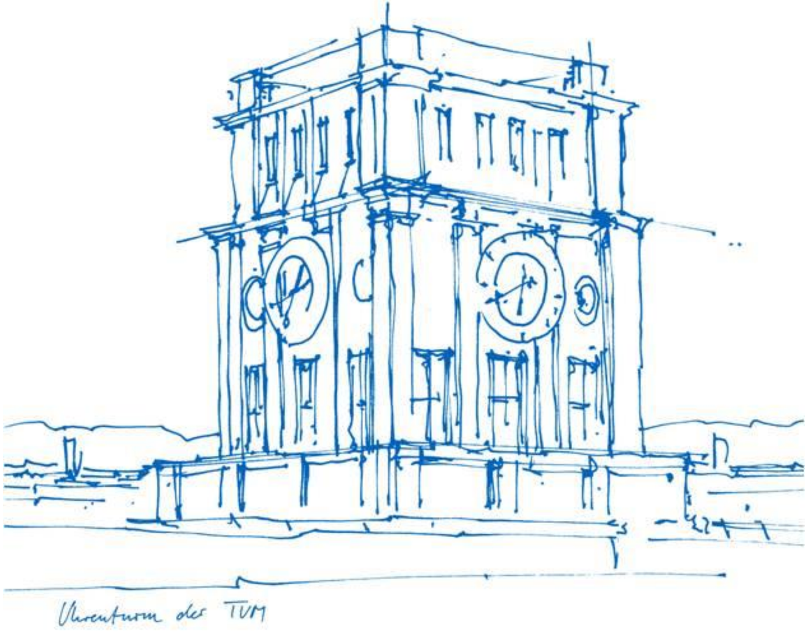
Volker Gabler

Tim Stahl, Gerold Huber, Ozgur Oguz, Dirk Wollherr

Chair of Automatic Control Engineering

Technical University of Munich

ICRA, Singapur, 31. May 2017



Motivation and General Approach

Autonomous Robots in Human-Robot Collaboration (HRC)

Given are high-level actions [Nau+ 2004] e.g.

```
pick(robot, object1)
pick(robot, object2)
pick(human, object1)
pick(human, object3)
```

The goal is to adapt high-level action-selection such that

- agents' effort is minimized
- team-efficiency is maximized
- human safety is guaranteed

Contribution

Autonomous Decisions in HRC

State of the art: Adapt to human action without reflecting human adaptivity [Mainprice+ 2013; Hawkins+ 2014; Maeda+ 2014; Gombolay+2015]

Contribution: Evaluation of the complete action-space for all involved agents using **normal form games**

Applied Game Theory in HRC

State of the art: Application limited to differential game theory or simulations [Jarrassé+ 2012; Li+ 2015; Bahram+ 2015; Turnwald+ 2016]

Contribution: discrete online action selection in real HRC

General Approach: Adaptive Action-Selection as a Normal Form Game

Iterative Decision Process as a **Game:**

- Finite Game
- Rational Players
- Complete Information
- Non-Zero-Sum
- Non-Cooperative
- Normal Form

Basic Assumptions

Direct mapping of high-level action and estimated trajectory

Interaction heuristics rather than purely data-driven models

Applied Interaction Heuristics

- Task dependent reward r_k
- Native cost c_k^{nat}
- Interactive cost c_k^{inter}

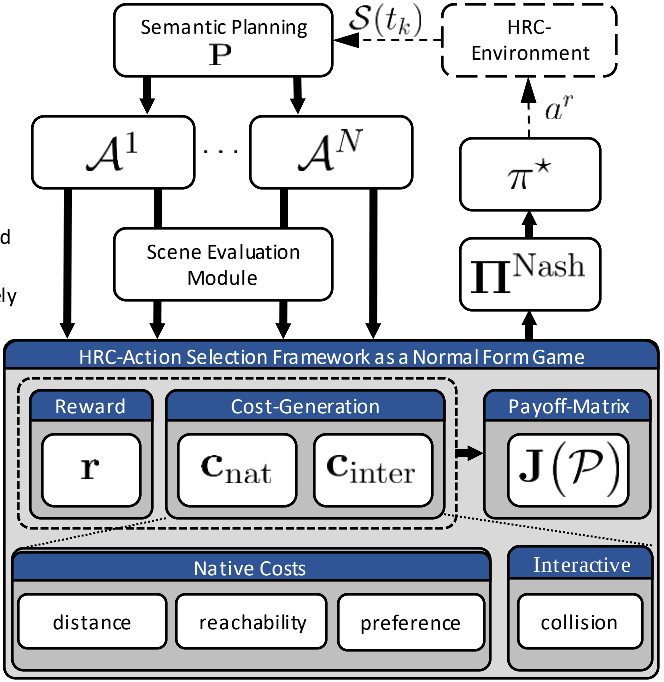
$$J_k(\boldsymbol{\pi}) = r_k - c_k^{\text{nat}}(a_k) - c_k^{\text{inter}}(\boldsymbol{\pi})$$
$$\text{with } a_k \in \boldsymbol{\pi}$$

Retrieve Nash-Equilibrium/-a

$$\boldsymbol{\pi}^\star \leftarrow \underset{\boldsymbol{\pi}^k \in \boldsymbol{\Pi}^{\text{Nash}}}{\operatorname{argmax}} \; u^i(\boldsymbol{\pi}^k)$$

$$\boldsymbol{\pi}^\star \leftarrow \underset{\boldsymbol{\pi}^k \in \boldsymbol{\Pi}^{\text{Nash}}}{\operatorname{argmax}} \; \sum_{i=0}^{N} u^i(\boldsymbol{\pi}^k)$$

$$\boldsymbol{\pi}^\star \leftarrow \underset{\boldsymbol{\pi}^k \in \boldsymbol{\Pi}^{\text{Nash}}}{\operatorname{isPareto}}(\boldsymbol{\pi}^k)$$



Cited Related Work

M. Bahram, A. Lawitzky, J. Friedrichs, M. Aeberhard and D. Wollherr. A Game Theoretic Approach to Replanning-aware Interactive Scene Prediction and Planning. In: IEEE Trans. Veh. Technol. 65.6 (2015), pp. 3981–3992.
M. C. Gombolay, R. A. Gutierrez, S. G. Clarke, D.GF.Sturl and J. A. Shah. Decision-Making Authority, Team Efficiency and Human Worker Satisfaction in Mixed Human-Robot Teams. In: Autonomous Robots (2015).
K. P. Hawkins, S. Bansal, N. N. Vo and A. F. Bobick. Anticipating human actions for collaboration in the presence of task and sensor uncertainty. In: ICRA. 2014, pp. 2215–2222.
N. Jarrass´e, Th. Charalambous and E. Burdet. A Framework to Describe, Analyze and Generate Interactive Motor Behaviors. In: PLoS ONE 7.11 (2012).
Y. Li, K. P. Tee, W. L. Chan, R. Yan, Y. Chua and D. K. Limbu. Role Adaptation of Human and Robot in Collaborative Tasks. In: ICRA. 2015, pp. 5602–5607.
G. Maeda, M. Ewerton, R. Lioutikov, H. B. Amor, J. Peters and G. Neumann. Learning interaction for collaborative tasks with probabilistic movement primitives. In: IEEE-RAS. 2014, pp. 527–534.
J. Mainprice and D. Berenson. Human-robot collaborative manipulation planning using early prediction of human motion. In: IROS. 2013, pp. 299–306.
D. Nau, M. Ghallab and P. Traverso. Automated Planning: Theory & Practice. Morgan Kaufmann Publishers Inc., 2004, p. 229. isbn: 1558608567.
A Turnwald, D Althoff, D Wollherr and M Buss. Understanding Human Avoidance Behavior: Interaction-Aware Decision Making Based on Game Theory. In: I. J. of Social Robotics 8.2 (2016), pp. 331–351.

Conclusion

Design of a normal form game decision framework

Online application of proposed framework

Confirmed three hypotheses in extensive user-study

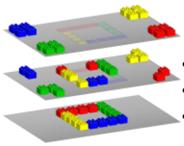
Future Work

Extension to multi-agent systems

Comparison with latest state-of-the-art on complex scenarios

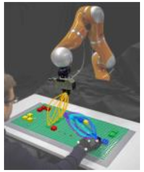
Experimental Insights I – General Overview, Results and Video

Experimental Setup



Baseline comparison to a fixed action policy

- Cooperative pick- and place assembly
- $n = 30$ participants
- All participants build same task under different policies applied



Policies Compared

1. *Fixed* - Non-reactive policy in which the robot is simply following.
2. *Line* - Proposed Framework with straight line human motion prediction
3. *Spline* - Proposed Framework with minimum jerk human motion prediction

Hypotheses

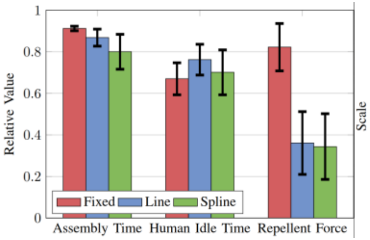
- **H1** - Participants prefer the robot's action-selection when working in the Spline or Line mode over the decisions in Fixed mode.
- **H2** - The decisions of the robot increase the safety for the human in Spline or Line mode, compared to the Fixed mode.
- **H3** - The robot's decisions adapt to the human and therefore decrease the overall completion time in the Spline or Line mode, compared to the Fixed mode.

Experimental Measurements

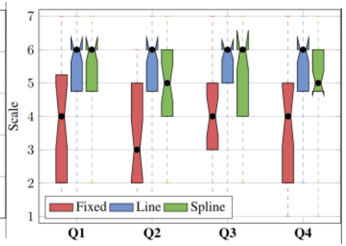
- Subjective questionnaire (**H1**) using a 7-point Likert scale
- Potential field based compliance control around robot end-effector to measure repellent force as a safety measure (**H2**)
- Overall completion time from first robot motion to assembly of last object(**H3**)

Experimental Results

Empirical Measurements



Subjective Feedback



How would you grade the ...

Q1 ... *collaboration with the robot?*
Q2 ... *robot as a helpful co-worker?*
Q3 ... *motion reaction of the robot?*
Q4 ... *action selection of the robot?*

Question	Overall Comparison	Line vs Spline	Line vs Random	Spline vs Random
Q1	**0.0016**	0.6670	**0.0034**	**0.0014**
Q2	**0.0004**	0.9324	**0.0009**	**0.0005**
Q3	**0.0032**	0.9327	**0.0025**	**0.0052**
Q4	**0.0021**	0.7709	**0.0023**	**0.0030**

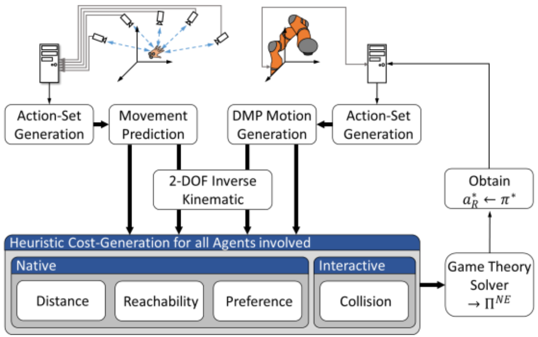
p-Values; statistically significant values shown in bold font

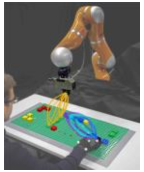
Results

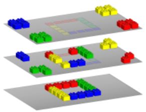
1. Subjective Evaluation proved statistic significant improvements for **H1.**
2. The reduced repellent force of the potential (safety) field confirms **H2.**
3. Except some minor overlap in the variance of the assembly time concerning *Line* vs. *Fixed* comparison, **H3** is confirmed.

Experimental Insights – Applied Heuristics and Video

Experimental Framework







Cost-Generation native components

Distance Based Costs

$$c^i_{\mathrm{nat}} = \alpha\, c^i_{\mathrm{travel}} + \beta\, c^i_{\mathrm{reach}} + \gamma\, c^i_{\mathrm{pref}}$$

$$c^i_{\mathrm{travel}} = \alpha_{\mathrm{L}} \int_{t_k}^{t_e} \dot{\mathbf{p}}^i_{\mathrm{s}}(t)\, dt + \alpha_{\mathrm{T}} \int_{t_k}^{t_e} 1\, dt + \alpha_{\mathrm{D}} ||\mathbf{p}_{\mathrm{e}} - \mathbf{p}^i_{\mathrm{s}}||$$

Reachability Based Costs

$$c^i_{\mathrm{reach}} = \begin{cases} \infty & \text{if } d > d_{\max} \\ \exp\left(w_{\mathrm{r}} \left(\frac{d - d_{\max}}{d_{\max}}\right)^2\right) & \text{else} \end{cases}$$

Preference Based Costs

$$c^i_{\mathrm{pref}} = \begin{cases} 0 & \text{if } ||\dot{\mathbf{p}}^i_{\mathrm{s}}|| < \epsilon_v \\ \varphi_{\mathrm{diff}}\, ||\dot{\mathbf{p}}^i_{\mathrm{s}}|| & \text{else} \end{cases} \qquad \varphi_{\mathrm{diff}} = \left| \cos^{-1}\left(\frac{\langle \mathbf{d}_{\mathrm{dest}}, \dot{\mathbf{p}}^i_{\mathrm{s}}(t) \rangle}{|\mathbf{d}_{\mathrm{dest}}|\, |\dot{\mathbf{p}}^i_{\mathrm{s}}(t)|} \right) \right|$$

Cost-Generation interactive component

Collision risk evaluation given two trajectories in x-y-plane

$$c^{i,j}_{\mathrm{inter,k}} = \begin{cases} c_{\mathrm{col}} & \text{if } d_{\mathrm{t}} < d_{\mathrm{cmin}} \\ \Psi(d_{\mathrm{t,k}}) & \text{if } d_{\mathrm{cmin}} < d_{\mathrm{t,k}} < d_{\mathrm{cmax}} \\ 0 & \text{else,} \end{cases}$$

$$c^{i,j}_{\mathrm{inter,avg}} = \frac{\sum_{k=1}^{N_{\mathrm{smpl}}} w_k^{\mathrm{temp}}\, c^{i,j}_{\mathrm{inter,k}}}{\sum_{k=1}^{N_{\mathrm{smpl}}} w_k^{\mathrm{temp}}}, c^{i,j}_{\mathrm{inter}} = \max(c^{i,j}_{\mathrm{inter,avg}}, c^{i,j}_{\mathrm{inter,max}}),$$

$$c^{i,j}_{\mathrm{inter,max}} = \max_{k \in \{1,2,\ldots,N_{\mathrm{smpl}}\}} \left(w_k^{\mathrm{temp}}\, c^{i,j}_{\mathrm{inter,k}} \right).$$

Robot Motion Control

Use a database of prelearned Dynamic Movement Primitives (DMPs):

$$\tau \ddot{\mathbf{y}} = \alpha_z(\beta_z(\mathbf{g} - \mathbf{y}) - \dot{\mathbf{y}}) + \mathbf{f}, \in \mathbb{R}^3,$$

Robotic Action set therefore given by