Transferring Impedance Control Strategies Between Heterogeneous Systems via Apprenticeship Learning

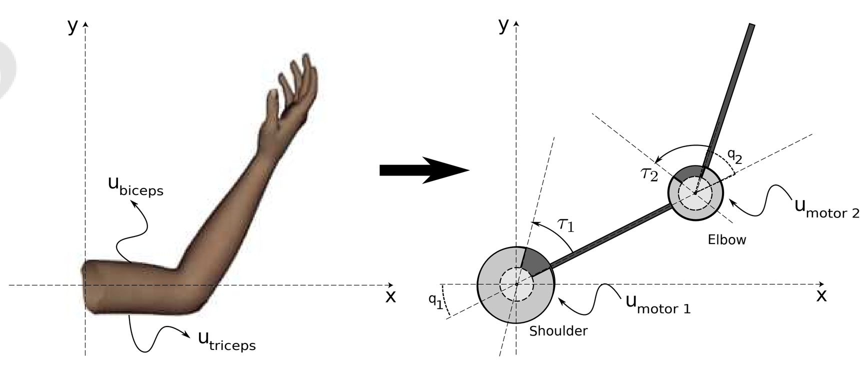
Matthew Howard, Djordje Mitrovic, Sethu Vijayakumar – *School of Informatics, The University of Edinburgh, UK*

1 – Introduction

Human behaviour, especially that related to impedance control, is highly adapted and optimised with respect to the complex, non-linear, noisy dynamics of the **musculo-skeletal system**.

Robotic systems have numerous designs, some biomorphic but are **built from different materials** and **have different actuation mechansims** $\rightarrow$ **very different dynamics**.

Schemes for **transferring dynamic behaviour** from human to robot **must** take into account these differences.



Our approach: **Inverse optimal control** to learn the **task objectives** in a way that is *independent of the dynamics*.

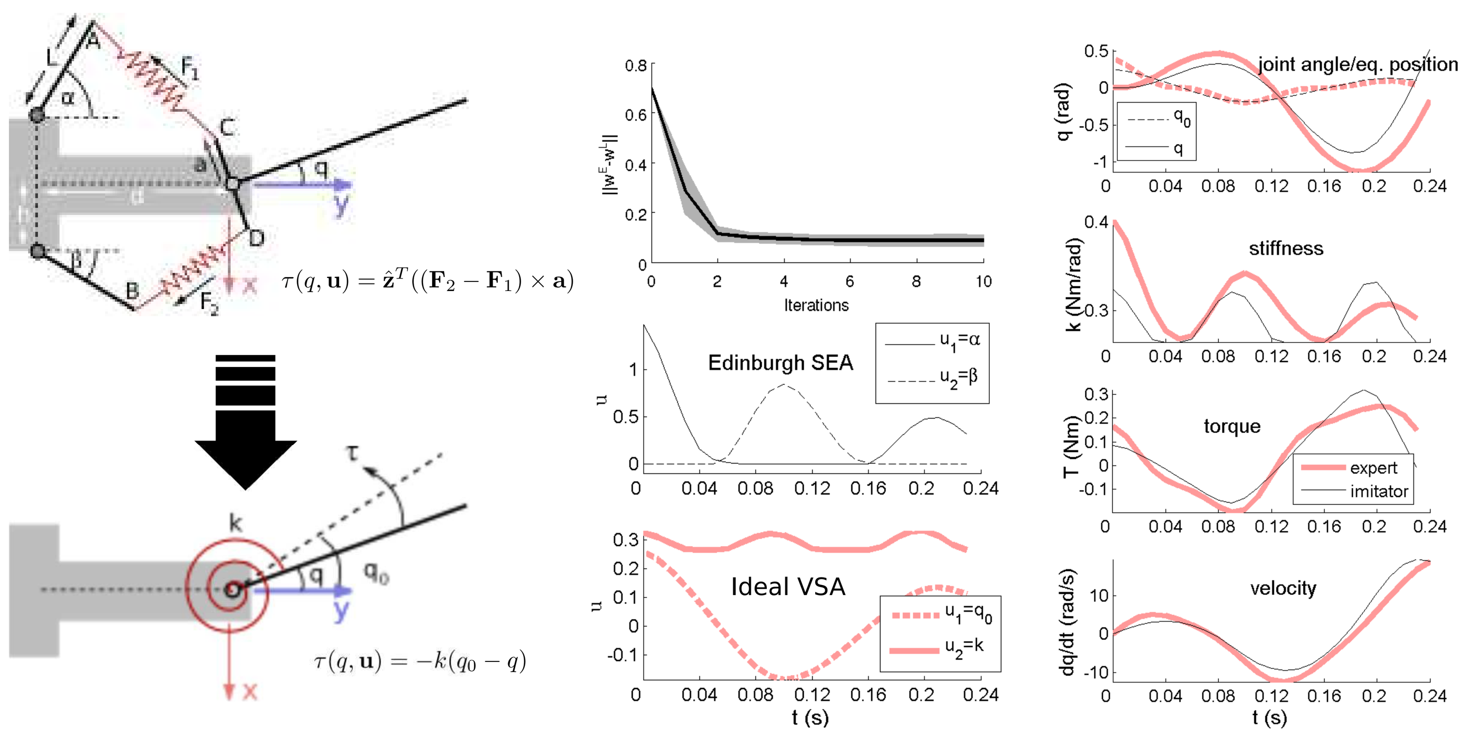
- Learnt objective gives **efficient representation** of the behaviour $\rightarrow$ **generalisation to systems with arbitrary dynamics**
- Behaviour transfer with the learnt objective function can even result in the imitator **surpassing the performance of the demonstrator**.

3 – Simulation Studies

a. Single-joint Systems

Transfer between single-joint variable stiffness actuators (VSAs) with **heterogeneous actuation mechanisms**. Demonstrations consist of optimal hitting movements for the **Edinburgh SEA** (a non-linear, antagonistic VSA).

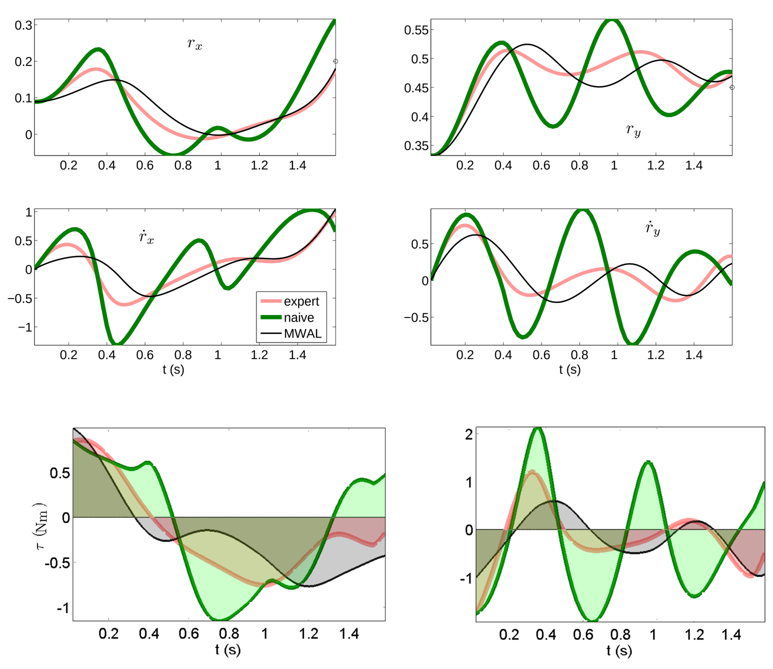
Objective function model: $\quad {}^eJ = w_1(q(T) - q^*)^2 - w_2\dot{q}(T) + \int_0^T w_3\tau^2\,dt$



Behaviour is transferred to an **'ideal' VSA** (with decoupled stiffness and position control). **Totally different** command sequence is found, but a **similar impedance strategy emerges.**

b. Multi-joint Systems

Transfer from **human arm** (6-muscle, 2-joint model) to **variable stiffness robot** (2-joint, decoupled stiffness/position control).



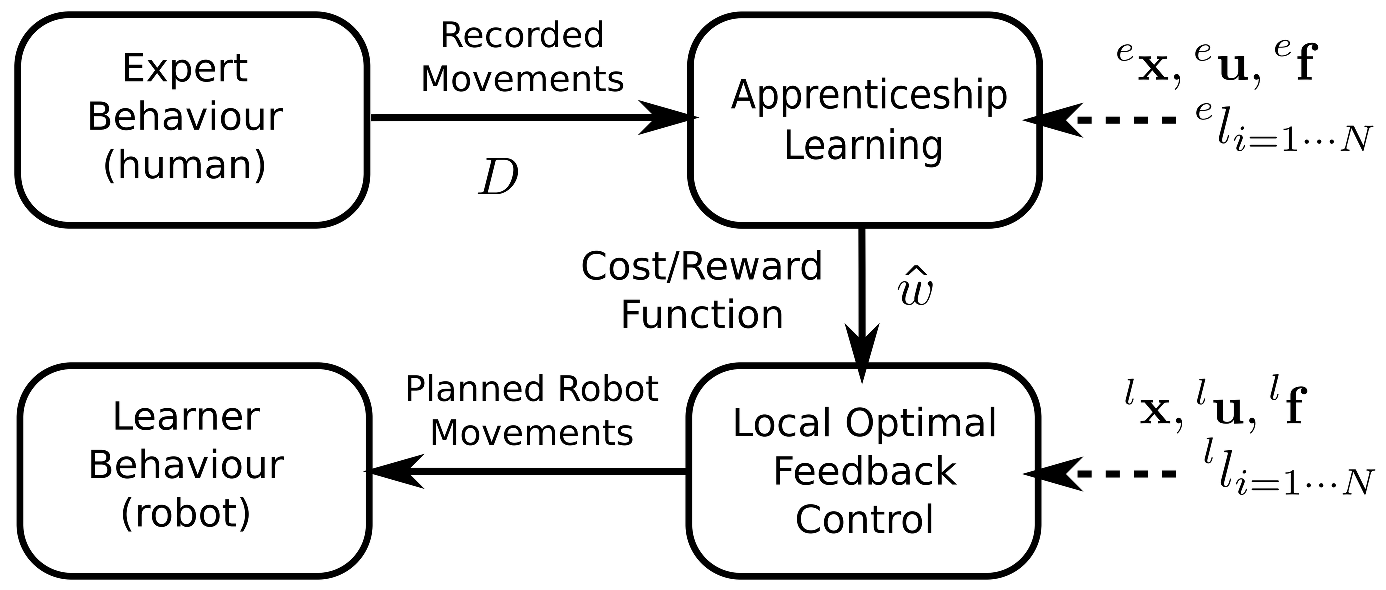
- **Punching task**

$$ {}^eJ = w_1\|\mathbf{r}(T) - \mathbf{r}^*\|_2^2 - w_2\dot{r}_x(T) + \int_0^T w_3\|\boldsymbol{\tau}\|_2^2\,dt $$

- **Compare** (i) **direct imitation of stiffness**, and (ii) **apprenticeship learning**.

Apprenticeship learning takes into account **difference in damping characteristics** $\rightarrow$ stable, efficient movement with **comparable performance to human demonstrator**.

2 – Method



1. Collect demonstrations (trajectories) $\{({}^e\mathbf{x}_0^k, {}^e\mathbf{u}_0^k), \cdots, ({}^e\mathbf{x}_T^k, {}^e\mathbf{u}_T^k)\}_{k=0}^K$. These are assumed optimal with respect to expert's

Dynamics	**Objective**
${}^e\dot{\mathbf{x}} = {}^e\mathbf{f}({}^e\mathbf{x}, {}^e\mathbf{u}) \in \mathbb{R}^n$ $= {}^e\mathbf{g}({}^e\mathbf{x}, {}^e\boldsymbol{\tau}({}^e\mathbf{x}, {}^e\mathbf{u}))$	${}^eJ = {}^eh({}^e\mathbf{x}(T)) + \int_0^T {}^el({}^e\mathbf{x}, {}^e\mathbf{u}, t)\,dt$

2. Multiplicative Weights Apprenticeship Learning (Syed et al., 2008) to **learn the objective function parameters**.

Objective function model:

$$ {}^eJ = \sum_{i=1}^{n_T} w_i\,{}^eh_i({}^e\mathbf{x}(T)) + \int_0^T \sum_{i=n_T}^{N} w_i\,{}^el_i({}^e\mathbf{x}, {}^e\mathbf{u}, t)\,dt \quad \in \mathbb{R} $$

3. Match expert objective function to **correspondent** learner objectives

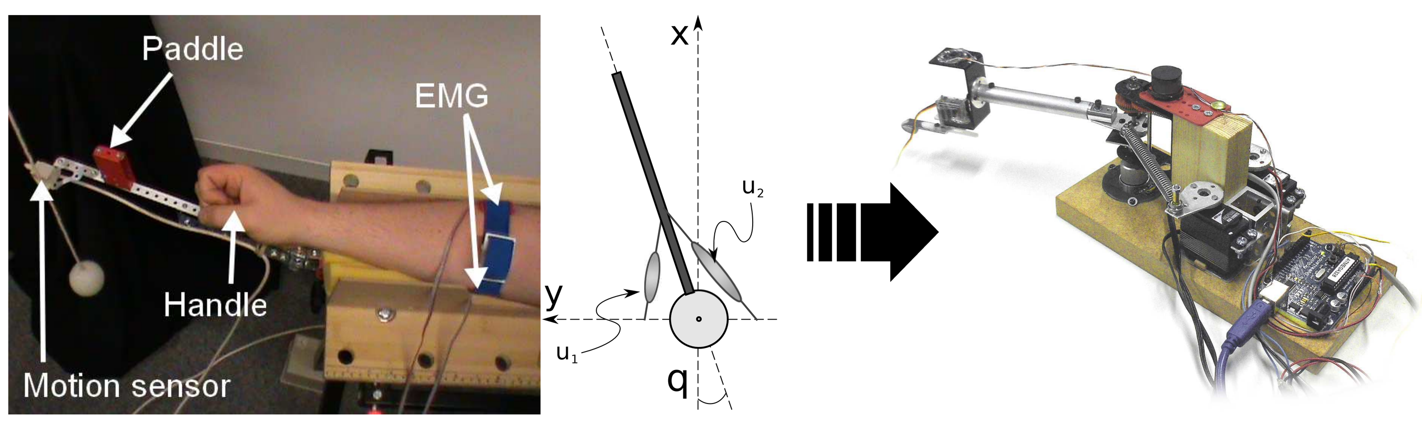
$$ {}^eh_i({}^e\mathbf{x}(T)) \longleftrightarrow {}^lh_i({}^l\mathbf{x}(T)) \qquad {}^el_i({}^e\mathbf{x}, {}^e\mathbf{u}, t) \longleftrightarrow {}^ll_i({}^l\mathbf{x}, {}^l\mathbf{u}, t) $$

4. Local Optimal Feedback Control (ILQG, Todorov and Li, 2005) to optimise **learnt objective function** under **imitator dynamics**.

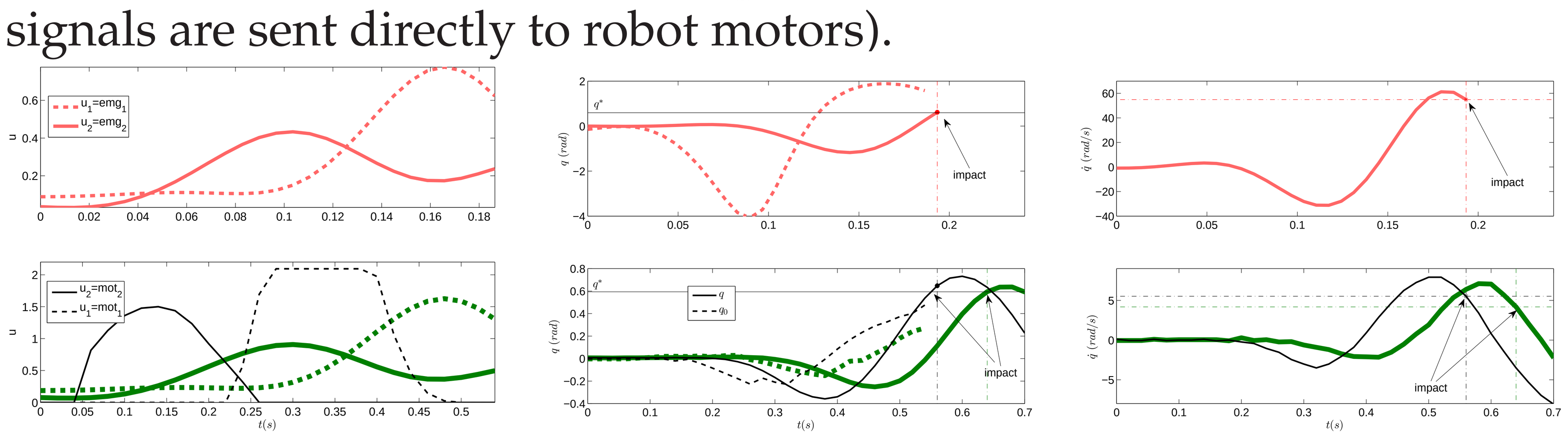
Optimise ${}^lJ = \sum \hat{w}_i\,{}^lh_i({}^l\mathbf{x}(T)) + \int_0^T \sum \hat{w}_i\,{}^ll_i({}^l\mathbf{x}, {}^l\mathbf{u}, t)\,dt$ under ${}^l\dot{\mathbf{x}} = {}^l\mathbf{f}({}^l\mathbf{x}, {}^l\mathbf{u})$.

4 – Human Experiment

Transfer from human demonstrations. Human is asked to 'hit ball as hard as possible'. State (wrist angle, angular velocity) and actions (antagonistic muscle activations measured with suface EMG) are recorded during movement. We transfer hitting to the **Edinburgh SEA hardware**.



Apprenticeship learning results in hitting behaviour adapted to the hardware - ball is **hit harder** compared to **direct imitation** (where EMG signals are sent directly to robot motors).



5 – Conclusion

Transfer of impedance behaviour between heterogeneous systems requires **differences in the dynamics** to be taken into account. **Apprenticeship learning** offers a principled approach to transfer behaviour at the level of the **task objectives** in a way **independent of the dynamics**. Experiments have shown an **inverse optimal control** approach results in improved performance with respect to the **demonstrator's task objectives** compared to standard imitation learning approaches.

References

Todorov & Li. *A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems.* Proceedings of the American Control Conference, 2005.
Syed, Bowling, & Schapire. *Apprenticeship learning using linear programming.* ICML, 2008.

Acknowledgements: This research is supported by the EU FP7 project STIFF.
Contact Author: Matthew Howard (matthew.howard@ed.ac.uk) for further details on this work.