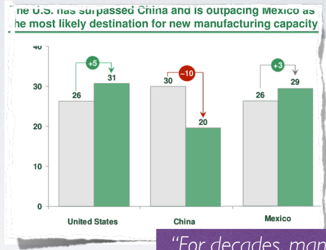
SEMINARS IN ARTIFICIAL INTELLIGENCE

HUMAN ROBOT COLLABORATION

Marc Hanheide

WHAT'S THE POINT?

- Automation with robots is on the rise
- "re-shoring" is a big deal these days



"For decades, manufacturers have had very few cost-effective options for handling low volume, high mix production jobs. No longer. Meet Baxter – the safe, flexible, affordable alternative to outsourced labor and fixed automation. Leading companies across North America have already integrated Baxter into their workforce, and gained a competitive advantage for their business in the process."

STRATEGIC RESEARCH AGENDA IN EUROPE

SRA

Robotics has the potential to transform many areas of business from manufacturing to healthcare. No more so than in areas where robots and people closely interact and **collaborate**. Robots are already serving as life-saving surgical tools, smart rehabilitation trainers, as well as reliable movers in all kinds of logistics scenarios; their role, impact and **interaction with people will only grow**.

Manufacturing

Technology

- Accurate indoor positioning systems for mobile manipulators, particularly in dynamic environments.
- Sensor based safety systems to enhance human robot interaction.
- Higher levels of realism in system modelling to speed application development.
- Reactive planning and control able to operate a robot safely in real industrial environments.

SRA

Robotics has the potential to transform many areas of business from manufacturing to healthcare. No more so than in areas where robots and people closely interact and **collaborate**. Robots are already serving as life-saving surgical tools, smart rehabilitation trainers, as well as reliable movers in all kinds of logistics scenarios; their role, impact and **interaction with people will only grow**.

UNIVERSITY OF LINCOLN

SAPIENZA UNIVERSITÀ DI ROMA

L-CAS

Manufacturing

Technology

- Accurate indoor positioning systems for mobile manipulators, particularly in dynamic environments.
- Sensor based safety systems to enhance human robot interaction.
- Higher levels of realism in system modelling to speed application development.
- Reactive planning and control able to operate a robot safely in real industrial environments.

Healthcare

Technology

- Improved teleoperation and physical interaction
- Miniaturised mechanical systems and sensing
- Multiple degree of freedom tactile feedback
- Inherently safe systems
- Monitoring of patient condition and improved data interpretation during procedures

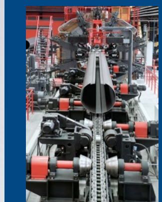
Robotics has the potential to transform many areas of business from manufacturing to healthcare. No more so than in areas where robots and people closely interact and **collaborate**. Robots are already serving as life-saving surgical tools, smart rehabilitation trainers, as well as reliable movers in all kinds of logistics scenarios; their role, impact and **interaction with people will only grow**.

UNIVERSITY OF LINCOLN

SAPIENZA UNIVERSITÀ DI ROMA

L-CAS

Manufacturing



Technology

- Accurate indoor positioning systems for mobile manipulators, particularly in dynamic environments.
- Sensor based safety systems to enhance human robot interaction.
- Higher levels of realism in system modelling to speed application development.
- Reactive planning and control able to operate a robot safely in real industrial environments.

Healthcare

Technology

- Improved teleoperation physical interaction
- Miniaturised mechanica systems and sensing
- Multiple degree of free tile feedback
- Inherently safe systems
- Monitoring of patient c and improved data inte tion during procedures

Logistics and Tran

Technology

- Interaction technology.
- Compliant mechanical systems.
- 3D environment interpretation.
- Task planning and optimisation.

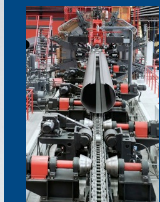
Robotics has the potential to transform many areas of business from manufacturing to healthcare. No more so than in areas where robots and people closely interact and **collaborate**. Robots are already serving as life-saving surgical tools, smart rehabilitation trainers, as well as reliable movers in all kinds of logistics scenarios; their role, impact and **interaction with people will only grow**.

UNIVERSITY OF LINCOLN

SAPIENZA UNIVERSITÀ DI ROMA

L-CAS

Manufacturing



Technology

- Accurate indoor positioning systems for mobile manipulators, particularly in dynamic environments.
- Sensor based safety systems to enhance human robot interaction.
- Higher levels of realism in system modelling to speed application development.
- Reactive planning and control able to operate a robot safely in real industrial environments.

Healthcare

Technology

- Improved teleoperation and physical interaction
- Miniaturised mechanical systems and sensing
- Multiple degree of freedom tactile feedback
- Inherently safe systems
- Monitoring of patient condition and improved data interpretation during procedures

Logistics and Tran

Technology

- Interaction technology.
- Compliant mechanical systems.
- 3D environment interpretation.
- Task planning and optimisation.

Robotics has the potential to transform many areas of business from manufacturing to healthcare. No more so than in areas where robots and people closely interact and **collaborate**. Robots are already serving as life-saving surgical tools, smart rehabilitation trainers, as well as reliable movers in all kinds of logistics scenarios; their role, impact and **interaction with people will only grow**.

UNIVERSITY OF LINCOLN

SAPIENZA UNIVERSITÀ DI ROMA

L-CAS

Manufacturing

Technology

- Accurate indoor positioning systems for mobile manipulators, particularly in dynamic environments.
- Sensor based safety systems to en-

Healthcare

Technology

- Improved teleoperation and physical interaction
- Miniaturised mechanical systems and sensing

Logistics and Tran

Technology

- Interaction technology.
- Compliant mechanical systems.
- 3D environment interpretation.

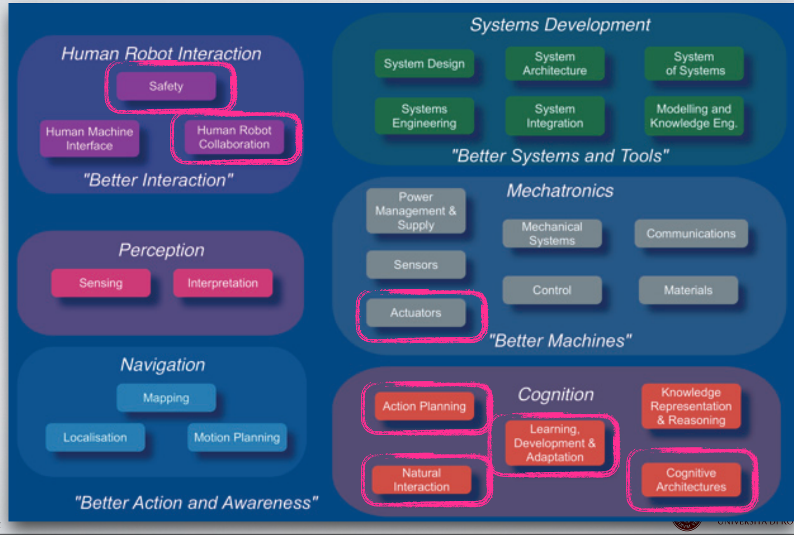
Robotics has the potential to transform many areas of business from manufacturing to healthcare. No more so than in areas where robots and people closely interact and **collaborate**. Robots are already serving as life-saving surgical tools, smart rehabilitation trainers, as well as reliable movers in all kinds of logistics scenarios; their role, impact and **interaction with people will only grow**.

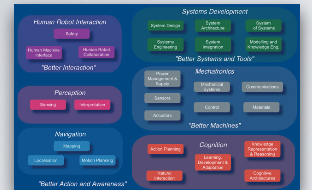
To Improve Human Robot Interaction. Better Collaboration:

Advanced robots will directly interact with users, physically and cognitively, these technologies will define our experience of robots and shape emerging applications.

SRA

To Improve Human Robot Interaction. Better Collaboration: Advanced robots will directly interact with users, physically and cognitively, these technologies will define our experience of robots and shape emerging applications.





SRA

To Improve Human Robot Interaction. Better Collaboration: Advanced robots will directly interact with users, physically and cognitively, these technologies will define our experience of robots and shape emerging applications.

Human Robot Interaction: "Better Interaction"

Advanced robots will increasingly interact with people. Physical and cognitive interactions are are at the core of many new areas of robot application. Robots will become tools used by people. Interaction will shift from computer like interfaces to being direct and physical. Robots collaborating and co-working with people both at home and in the workplace will become the norm. These technologies are fundamental to this step change. This increased physical interaction demands higher levels of safety and dependability. These technologies will provide the basis for building safe usable interactive machines.

Human Machine Interface

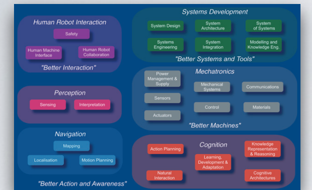
Robots will increasingly interact with people. This interaction will be essential to the acceptance and integration of robots into our everyday lives. It might be through buttons and a screen, or through physical interaction and gestures. Interaction will move from computer like interfaces to ones based on intuitive interpretation of a user's intentions.

State of the Art

Touch screen interaction is commonplace, and limited gesture recognition is now available in commercial products. Emotion recognition based on enhanced face recognition is available in the research laboratory, gaze tracking and speech recognition in quiet environments are now commonplace.

2020 Target

To develop instruct-able interfaces. To develop physically interactive interfaces for collaborative working. To develop interfaces that can assess the emotional and cognitive state of the user and respond appropriately. To develop standardised interfaces for autonomous appliances.



SRA

Human Robot Interaction: "Better Interaction"

Advanced robots will increasingly interact with people. Physical and cognitive interactions are are at the core of many new areas of robot application. Robots will become tools used by people. Interaction will shift from computer like interfaces to being direct and physical. Robots collaborating and co-working with people both at home and in the workplace will become the norm. These technologies are fundamental to this step change. This increased physical interaction demands higher levels of safety and dependability. These technologies will provide the basis for building safe usable interactive machines.

Human Machine Interface

Robots will increasingly interact with people. This interaction will be essential to the acceptance and integration of robots into our everyday lives. It might be through buttons and a screen, or through physical interaction and gestures. Interaction will move from computer like interfaces to ones based on intuitive interpretation of a user's intentions.

State of the Art

Touch screen interaction is commonplace, and limited gesture recognition is now available in commercial products. Emotion recognition based on enhanced face recognition is available in the research laboratory, gaze tracking and speech recognition in quiet environments are now commonplace.

2020 Target

To develop instruct-able interfaces. To develop physically interactive interfaces for collaborative working. To develop interfaces that can assess the emotional and cognitive state of the user and respond appropriately. To develop standardised interfaces for autonomous appliances.

Human Robot Collaboration

The ability to physically interact with people a fundamental requirement for the next generation of robots. Many potential applications depend on developing safe robots that can provide autonomous, intuitive, physical interaction. To achieve this, human collaboration and safety criteria need to be placed at the centre of the design process. Intuitive physical interaction between humans and robot systems need to be addressed in an interconnected manner. The goal of this technology is to enable the close, safe and dependable physical interaction between people and robots in a shared workspace.

State of the Art

Simple interactions are commonly used, collaboration has been the subject of extensive research. Compliant systems are well understood and commercial products are starting to emerge. Safety standards are starting to emerge.

2020 Target

To develop low cost safe dependable systems able to react and interact with people. To understand the bio-mechanics of human injury and motion. To track, understand and predict human motion, in real-time, in specific environments. To integrate cognition technologies into human robot collaboration. To develop tools for safety validation. To develop safety standards. To develop multi-modal collaboration.

Human Robot Collaboration

The ability to physically interact with people a fundamental requirement for the next generation of robots. Many potential applications depend on developing safe robots that can provide autonomous, intuitive, physical interaction. To achieve this, human collaboration and safety criteria need to be placed at the centre of the design process. Intuitive physical interaction between humans and robot systems need to be addressed in an interconnected manner. The goal of this technology is to enable the close, safe and dependable physical interaction between people and robots in a shared workspace.

State of the Art

Simple interactions are commonly used, collaboration has been the subject of extensive research. Compliant systems are well understood and commercial products are starting to emerge. Safety standards are starting to emerge.

2020 Target

To develop low cost safe dependable systems able to react and interact with people. To understand the bio-mechanics of human injury and motion. To track, understand and predict human motion, in real-time, in specific environments. To integrate cognition technologies into human robot collaboration. To develop tools for safety validation. To develop safety standards. To develop multi-modal collaboration.

SRA

Human Robot Interaction: "Better Interaction"

Advanced robots will increasingly interact with people. Physical and cognitive interactions are are at the core of many new areas of robot application. Robots will become tools used by people. Interaction will shift from computer like interfaces to being direct and physical. Robots collaborating and co-working with people both at home and in the workplace will become the norm. These technologies are fundamental to this step change. This increased physical interaction demands higher levels of safety and dependability. These technologies will provide the basis for building safe usable interactive machines.

Human Machine Interface

Robots will increasingly interact with people. This interaction will be essential to the acceptance and integration of robots into our everyday lives. It might be through buttons and a screen, or through physical interaction and gestures. Interaction will move from computer like interfaces to ones based on intuitive interpretation of a user's intentions.

State of the Art

Touch screen interaction is commonplace, and limited gesture recognition is now available in commercial products. Emotion recognition based on enhanced face recognition is available in the research laboratory, gaze tracking and speech recognition in quiet environments are now commonplace.

2020 Target

To develop instruct-able interfaces. To develop physically interactive interfaces for collaborative working. To develop interfaces that can assess the emotional and cognitive state of the user and respond appropriately. To develop standardised interfaces for autonomous appliances.

Safety

Robots must be safe to use. Safety is a critically important aspect of robot operation, both in an industrial setting and when robots are interacting closely with people. Safety must be designed into a system, and tested according to well defined standards.

State of the Art

Safety is widely implemented through the exclusion of people from operating environments. In most cases physical barriers ensure a safe operating environment. Safety critical software development processes are used in some areas of robotics. The lack of standardised methods, development tools and verification criteria currently limit reusability.

2020 Target

To develop robust safety based design processes including inherent physical robot safety. Development of standards and methods to verify and certify safety in human robot collaboration. To create software development methods and tools which support the creation of solutions under safety constraints. To create software based safety systems providing dependable failure mode detection and isolation. To develop safety systems for multiple distributed robot systems. To develop predictive systems to assess the safety of human interaction.

Human Robot Collaboration

The ability to physically interact with people a fundamental requirement for the next generation of robots. Many potential applications depend on developing safe robots that can provide autonomous, intuitive, physical interaction. To achieve this, human collaboration and safety criteria need to be placed at the centre of the design process. Intuitive physical interaction between humans and robot systems need to be addressed in an interconnected manner. The goal of this technology is to enable the close, safe and dependable physical interaction between people and robots in a shared workspace.

State of the Art

Simple interactions are commonly used, collaboration has been the subject of extensive research. Compliant systems are well understood and commercial products are starting to emerge. Safety standards are starting to emerge.

2020 Target

To develop low cost safe dependable systems able to react and interact with people. To understand the bio-mechanics of human injury and motion. To track, understand and predict human motion, in real-time, in specific environments. To integrate cognition technologies into human robot collaboration. To develop tools for safety validation. To develop safety standards. To develop multi-modal collaboration.

SRA

Human Robot Interaction: "Better Interaction"

Advanced robots will increasingly interact with people. Physical and cognitive interactions are are at the core of many new areas of robot application. Robots will become tools used by people. Interaction will shift from computer like interfaces to being direct and physical. Robots collaborating and co-working with people both at home and in the workplace will become the norm. These technologies are fundamental to this step change. This increased physical interaction demands higher levels of safety and dependability. These technologies will provide the basis for building safe usable interactive machines.

Human Machine Interface

Robots will increasingly interact with people. This interaction will be essential to the acceptance and integration of robots into our everyday lives. It might be through buttons and a screen, or through physical interaction and gestures. Interaction will move from computer like interfaces to ones based on intuitive interpretation of a user's intentions.

State of the Art

Touch screen interaction is commonplace, and limited gesture recognition is now available in commercial products. Emotion recognition based on enhanced face recognition is available in the research laboratory, gaze tracking and speech recognition in quiet environments are now commonplace.

2020 Target

To develop instruct-able interfaces. To develop physically interactive interfaces for collaborative working. To develop interfaces that can assess the emotional and cognitive state of the user and respond appropriately. To develop standardised interfaces for autonomous appliances.

Safety

Robots must be safe to use. Safety is a … ts are interacting closely with people. Safety …

State of the Art

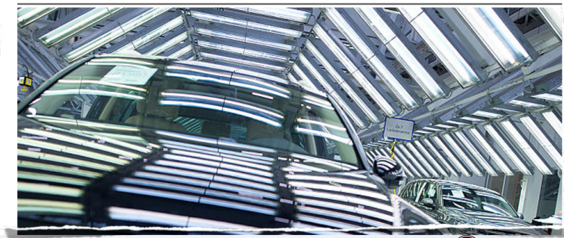
Safety is widely implemented throug… from operating environments. In mo… sure a safe operating environment. S… opment processes are used in some… of standardised methods, developm… teria currently limit reusability.

HOME » NEWS » WORLD NEWS » EUROPE » GERMANY

Robot kills man at Volkswagen plant in Germany

A 22-year-old worker was grabbed by the robot and crushed against a metal plate

7K 0 106 8K Email



ISO 10218: ROBOTS AND ROBOTIC DEVICES - SAFETY REQUIREMENTS FOR INDUSTRIAL ROBOTS

Hereinafter, the four different basic safety principles applicable in HRC are explained:

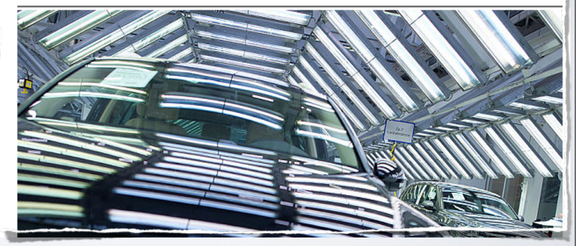
1. **Safety-rated monitored stop**
   Robot stops when operator enters the collaborative workspace and continues when the operator has left the collaborative workspace.
2. **Hand guiding**
   Robot movements are controlled by the operator.
3. **Speed and separation monitoring**
   Contact between operator and moving robot is prevented by the robot.
4. **Power and force limiting**
   Contact forces between operator and robot are technically limited to a safe level.

HOME » NEWS » WORLD NEWS » EUROPE » GERMANY

Robot kills man at Volkswagen plant in Germany

A 22-year-old worker was grabbed by the robot and crushed against a metal plate

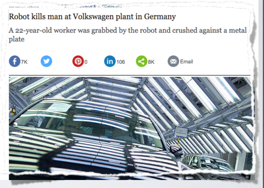
7K 0 106 8K Email



ISO 10218: ROBOTS AND ROBOTIC DEVICES - SAFETY REQUIREMENTS FOR INDUSTRIAL ROBOTS

Hereinafter, the four different basic safety principles applicable in HRC are explained:

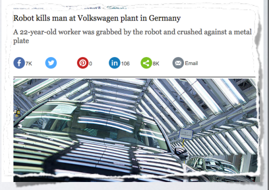
1. **Safety-rated monitored stop**
   Robot stops when operator enters the collaborative workspace and continues when the operator has left the collaborative workspace.
2. **Hand guiding**
   Robot movements are controlled by the operator.
3. **Speed and separation monitoring**
   Contact between operator and moving robot is prevented by the robot.
4. **Power and force limiting**
   Contact forces between operator and robot are technically limited to a safe level.



Conclusion

- Without risk assessment human-robot collaboration cannot take place.
- The overall application must always be considered (process, fixtures, gripper technology, robot), i.e. not only the robot.
- Safety functions must be implemented using suitable components in accordance to determined requirements.

ISO 10218: ROBOTS AND ROBOTIC DEVICES - SAFETY REQUIREMENTS FOR INDUSTRIAL ROBOTS

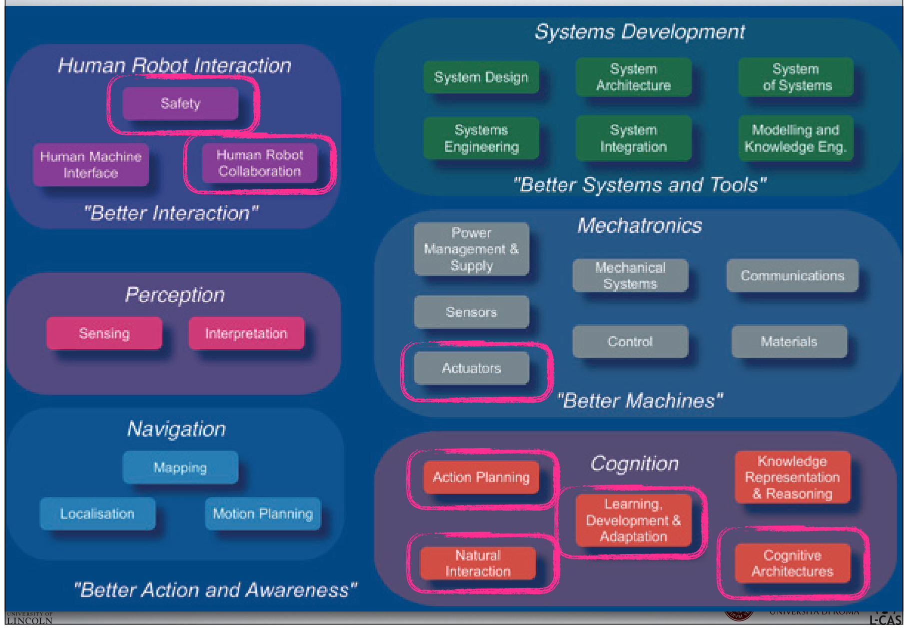


Conclusion

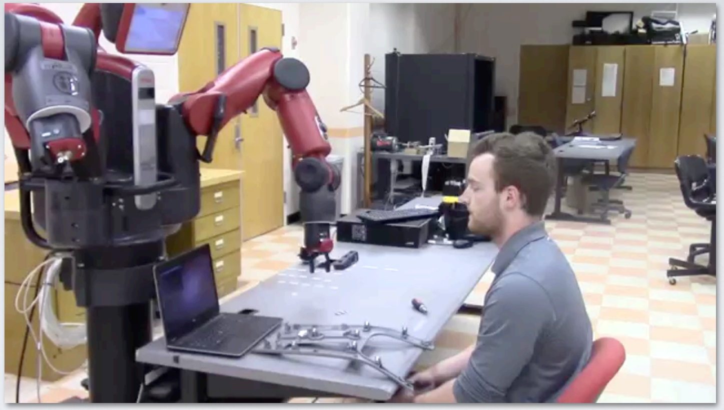
- Without risk assessment human-robot collaboration cannot take place.
- The overall application must always be considered (process, fixtures, gripper technology, robot), i.e. not only the robot.
- Safety functions must be implemented usin suitable components in accordance to deter mined requirements.

SRA

- **Cognitive human robot collaboration**
- Combines complex control of mechanical structures with ***interpretation*** of the collaborating partner's ***actions*** and their ***cognitive contexts*** with respect to the environment. Also includes ***mutual interaction*** with objects in the environment, for example the
  - shared lifting of an object,
  - providing assistance when standing up,
  - or the exchange and use of an object or tool.



ASPECTS OF HUMAN-ROBOT TEAMS



HUMAN ROBOT TEAMS

- Robots as **Teammates** (Hoffman & Breazeal, 2004):
  - implying a sense of **partnership** that occurs when agents work "jointly with" others rather than "acting upon" others
  - social adeptness on the robot's part
  - reason about our **intentions, beliefs, desires**, and **goals** so that it can perform the **right actions** at the appropriate time
  - **commitment** to doing their **own part**, commitment to the other in doing theirs, and commitment to the **success of the overall task**

SHARED/JOINT ACTIVITY

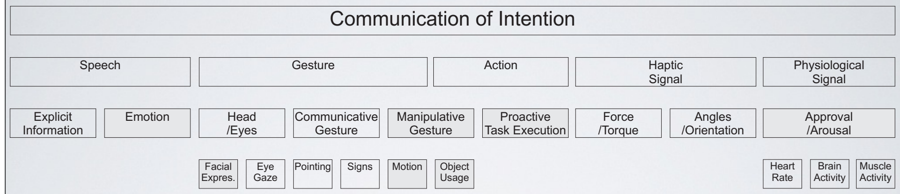
- Shared Cooperative Activity
  - mutual responsiveness
  - commitment to the joint activity
  - commitment to mutual support
- "meshing" singular sub-plans into a joint activity

"For example, if we were to move a table jointly through a doorway, your picking up one side of the table and starting to walk through the door does not make sense outside our joint activity. Even the sum of both our picking-up and moving actions would not amount to the shared activity without the existence of a collaborative plan that both of us are sharing (namely to move the table out the door). It seems that we both hold a joint intention, as well as individual intentions related to this joint intention."

JOINT INTENTION

- **mutual belief** about the state of the goal
- Teammates are **committed to inform** the team when they reach the conclusion that a goal is achievable, impossible, or irrelevant.
- efficient and robust collaboration scheme in a changing environment requires an **open channel of communication**

JOINT INTENTION



- **mutual belief** about the state of the goal
- Teammates are **committed to inform** the team when they reach the conclusion that a goal is achievable, impossible, or irrelevant.
- efficient and robust collaboration scheme in a changing environment requires an **open channel of communication**

COMMON GROUND

- central to collaboration is establishment / maintenance of **common ground:**
  - Clark: "*the sum of [...] mutual, common, or joint knowledge, beliefs, or suppositions*"
- with respect to the
  - objects of the task,
  - task state,
  - states of team members.
- **Joint closure** on a sub-activity is needed to advance a shared activity:
  - "participants in a joint action trying to establish the **mutual belief** that they have **succeeded well enough** for current purposes".
- Perspective Taking / Theory of Mind

GOALS

- Humans are biased to use an intention-based psychology to **interpret** other agent's actions ("*one cannot not communicate*")
- Humans interpret intentions and actions based on intended goals rather than specific activities or motion trajectories

*"A person opening a door, once by using her right hand and once by using her left hand, is considered to have performed the same action, regardless of the different **motion trajectories**. But this person is doing a distinctly different action if she is opening a tap, even though the motion might be very similar to the one used for opening the door."*

TURN TAKING

- Turn-taking is the fundamental way that humans organise interactions with each other
- In HRI, work on turn-taking needs to be approached from two directions:
  - awareness of the human's cue usage.: perception problem closely related to recognising contingency and engagement.
  - executing turn-taking cues in a socially intelligent manner.

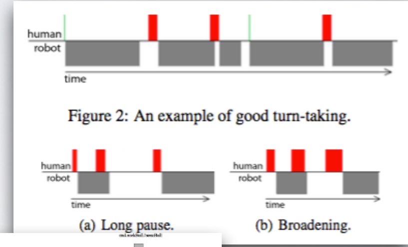
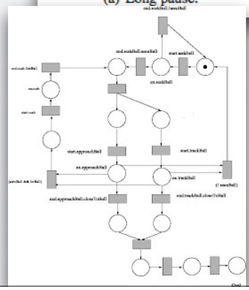
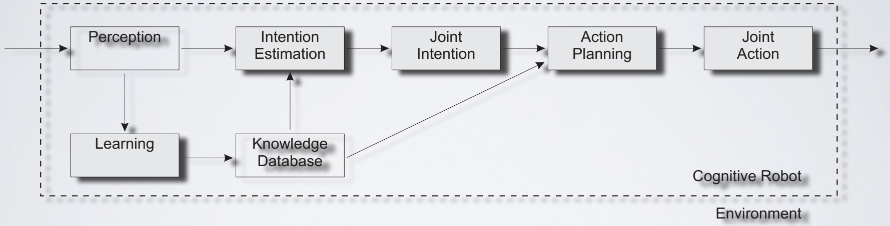


Figure 2: An example of good turn-taking.

(a) Long pause. (b) Broadening.



OVERVIEW



A more conceptual architecture "summarising" what has been said so far

5-1 St. Clair, A. & Mataric, M., 2011. Task coordination and assistive opportunity detection via social interaction in collaborative human-robot tasks. In Proc. Int. Conf. on Collaboration Technologies and Systems (CTS). IEEE, pp. 168–172. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5928682 [Accessed December 10, 2012].	Gabriele Angeletti	riccardo rinaldi

PLANNING TO COLLABORATE

PLANNING TO COLLABORATE

In planning it is typical to define a problem as a set of states with a set of possible associated actions.

These actions may cause a transition from one state to another. Ideally a transition is caused from the current state to a desired state which is derived from joint intention.

The general problem of planning is deciding upon a series of actions that will lead to reaching a goal, given the current state.

PLANNING TO COLLABORATE

In planning it is typical to define a problem as a set of states with a set of possible associated actions.

These actions may cause a transition from one state to another. Ideally a transition is caused from the current state to a desired state which is derived from joint intention.

The general problem of planning is deciding upon a series of actions that will lead to reaching a goal, given the current state.

Decision theoretic planning

- defines goal achievement through **maximising reward**.
- uses sequential decision models to decide on a **sequence of actions** to take in a series of states, rather than a one-time decision.
- sequential decision models that represent **uncertainty, time-dependence, history dependence, and partial observability**.
- per se **not suitable** for reasoning about **multiple partners** with different goals; multiple partners are combined into a model that presents the entire group as one single entity
- find a **policy** for only one partner acting in the presence of other partners, the other partners and their actions must be modelled as part of the state of the environment.

PLANNING TO COLLABORATE

In planning it is typical to define a problem as a set of states with a set of possible associated actions.

These actions may cause a transition from one state to another. Ideally a transition is caused from the current state to a desired state which is derived from joint intention.

The general problem of planning is deciding upon a series of actions that will lead to reaching a goal, given the current state.

multi-body planning

- the partners of a team can either be modelled as a **single body taking joint** actions or as a **game** where all the players have common payoffs.
- Complications may arise when not all partners **have the same knowledge** about the state of the environment or each others' abilities and it is impossible to build an accurate centralized model.
- In human-robot interaction, only the robotic partners can be controlled, plans and actions of the human peers are not guaranteed.

PLANNING TO COLLABORATE

In planning it is typical to define a problem as a set of states with a set of possible associated actions.

These actions may cause a transition from one state to another. Ideally a transition is caused from the current state to a desired state which is derived from joint intention.

The general problem of planning is deciding upon a series of actions that will lead to reaching a goal, given the current state.

total-order planning

- a plan is produced from sub plans **chronologically** from a starting state to an end state.
- It is defined as a search in the action space, as suitable actions are found chronologically for the sub problems.
- linearization of the sub plans happens during the whole planning process.

TOTAL ORDER PLANNING

Domain

Domain Examples

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;; 4 Op-blocks world
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

(define (domain BLOCKS)
  (:requirements :strips)
  (:predicates (on ?x ?y)
         (ontable ?x)
         (clear ?x)
         (handempty)
         (holding ?x)
         )

  (:action pick-up
        :parameters (?x)
        :precondition (and (clear ?x) (ontable ?x) (handempty
        :effect
        (and (not (ontable ?x))
         (not (clear ?x))
         (not (handempty))
         (holding ?x)))

  (:action put-down
        :parameters (?x)
```

Problem

Predefined Problems

```
(define
    (problem BLOCKS-3)
    (:domain BLOCKS)
    (:objects A B C )
    (:INIT
        (CLEAR B)
        (CLEAR C)
        (ONTABLE A)
        (ONTABLE B)
        (ON C A)
        (HANDEMPTY))
    (:goal (AND
        (ON A B) (ON B C)
        )
    )
)
```

Solve it

http://lcas.lincoln.ac.uk/fast-downward/

PLANNING TO COLLABORATE

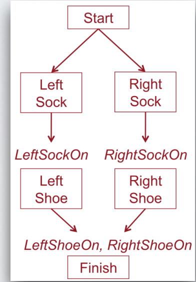
In planning it is typical to define a problem as a set of states with a set of possible associated actions.

These actions may cause a transition from one state to another. Ideally a transition is caused from the current state to a desired state which is derived from joint intention.

The general problem of planning is deciding upon a series of actions that will lead to reaching a goal, given the current state.

partial-order planning

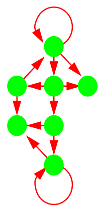
- sub problems are solved independently with sub plans, that are finally combined.
- search in the plan space, as first suitable plans are found for single **sub problems**.
- **Linearization of actions** happens in the **end** allowing for a higher flexibility.
- advantage is flexibility in the order in which an action plan is constructed.



MARKOV DECISION PROCESSES

A *Markov Decision Process* (MDP) model contains:

- A set of possible world states *S*
- A set of possible actions *A*
- A real valued reward function *R(s,a)*
- A description *T* of each action's effects in each state.



We assume the Markov Property: *the effects of an action taken in a state depend only on that state and not on the prior history.*

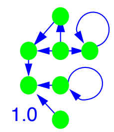
MARKOV DECISION PROCESSES

A *Markov Decision Process* (MDP) model contains:
- A set of possible world states *S*
- A set of possible actions *A*
- A real valued reward function *R(s,a)*
- A description *T* of each action's effects in each state.

We assume the Markov Property: *the effects of an action taken in a state depend only on that state and not on the prior history.*

Deterministic Actions:

- $T: \ S \times A \to S$ For each state and action we specify a new state.



Stochastic Actions:

- $T: \ S \times A \to Prob(S)$ For each state and action we specify a probability distribution over next states. Represents the distribution $P(s' | s, a)$.

MARKOV DECISION PROCESSES

A *Markov Decision Process* (MDP) model contains:

- A set of possible world states S
- A set of possible actions A
- A real valued reward function $R(s,a)$
- A description T of each action's effects in each state.

We assume the Markov Property: *the effects of an action taken in a state depend only on that state and not on the prior history.*

Deterministic Actions:

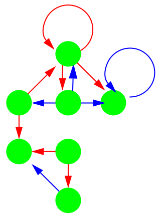
- $T: S \times A \rightarrow S$ For each state and action we specify a new state.

1.0

Stochastic Actions:

- $T: S \times A \rightarrow Prob(S)$ For each state and action we specify a probability distribution over next states. Represents the distribution $P(s' | s, a)$.

A *policy* π is a mapping from S to A



MARKOV DECISION PROCESSES

A *Markov Decision Process* (MDP) model contains:

- A set of possible world states S
- A set of possible actions A
- A real valued reward function $R(s,a)$
- A description T of each action's effects in each state.

We assume the Markov Property: *the effects of an action taken in a state depend only on that state and not on the prior history.*

Deterministic Actions:

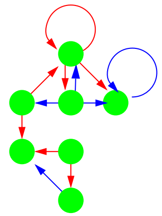
- $T: S \times A \rightarrow S$ For each state and action we specify a new state.

1.0

Stochastic Actions:

- $T: S \times A \rightarrow Prob(S)$ For each state and action we specify a probability distribution over next states. Represents the distribution $P(s' | s, a)$.

A *policy* π is a mapping from S to A



PARTIALLY OBSERVABLE MARKOV DECISION PROCESSES

System state can not always be determined

⇒ a Partially Observable MDP (POMDP)

- Action outcomes are not fully observable
- Add a set of observations O to the model
- Add an observation distribution $U(s,o)$ for each state
- Add an initial state distribution I

MARKOV DECISION PROCESSES

Nikolaidis, S. & Shah, J., 2013. Human-robot cross-training: Computational formulation, modeling and evaluation of a human team training strategy. In ACM/IEEE International Conference on Human-Robot Interaction. pp. 33–40.

Koppula, H.S., Jain, A. & Saxena, A., 2016. Anticipatory planning for human-robot teams. Springer Tracts in Advanced Robotics, 109(i), pp.453–470.

Koppula, H.S., Jain, A. & Saxena, A., 2016. Anticipatory planning for human-robot teams. Springer Tracts in Advanced Robotics, 109(i), pp.453–470.

which objects associated with which actions

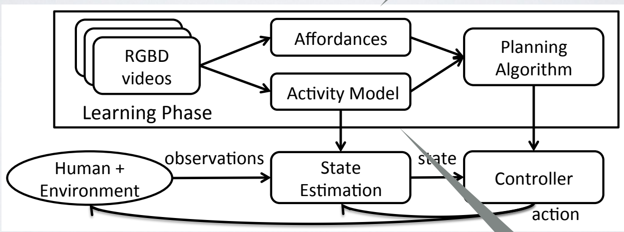


Fig. 2: **System Overview.**

human action "library"

Koppula, H.S., Jain, A. & Saxena, A., 2016. Anticipatory planning for human-robot teams. Springer Tracts in Advanced Robotics, 109(i), pp.453–470.

MDP

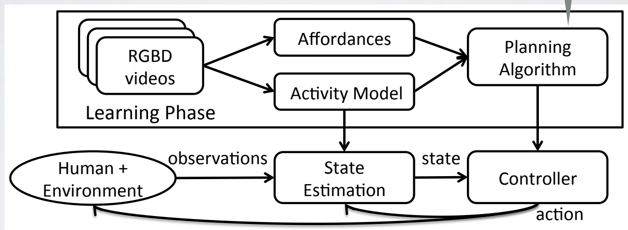


Fig. 2: **System Overview.**

Koppula, H.S., Jain, A. & Saxena, A., 2016. Anticipatory planning for human-robot teams. Springer Tracts in Advanced Robotics, 109(i), pp.453–470.

- State Space $\mathcal{S}$: Let $\mathbf{s_t} = \{s_t^1, ..., s_t^n\}$ denote the state of the environment, where s_t^i denotes the state of the i^{th} object at time t and n denotes the number of objects.
- Action Space $\mathcal{A}$: Let $a_t = \langle a_t^h, a_t^r \rangle$ denote the joint action at time t, where a_t^h and a_t^r denote the human and robot actions respectively.
- Robot's policy $\pi^r\colon \mathcal{S} \times \mathcal{A}^r \to [0,1]$, where $\mathcal{A}^r$ denotes the set of possible robot actions. $\pi^r(\mathbf{s}, a^r)$ specifies the probability of choosing action a^r in state $\mathbf{s}$.

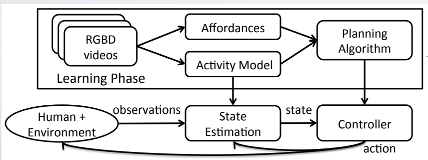


Fig. 2: **System Overview.**

MDP

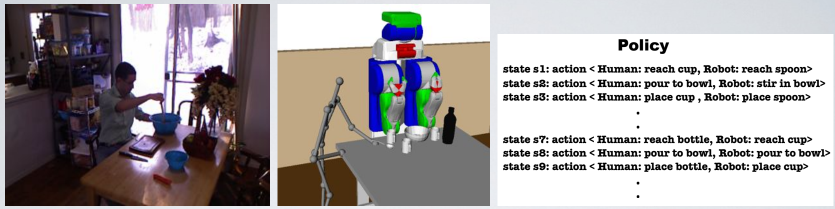


Fig. 3: **Collaborative planning by the robot.** In order to collaborate with the human on a *recipe following* task, the robot learns the activity model from RGB-D videos of human preparing a recipe (left), represents the environment via affordances and uses our planning algorithm (middle) to generate a policy for jointly performing the activity with the human (right).

- State Space $\mathcal{S}$: Let $\mathbf{s_t} = \{s_t^1, ..., s_t^n\}$ denote the state of the environment, where s_t^i denotes the state of the i^{th} object at time t and n denotes the number of objects.
- Action Space $\mathcal{A}$: Let $a_t = \langle a_t^h, a_t^r \rangle$ denote the joint action at time t, where a_t^h and a_t^r denote the human and robot actions respectively.
- Robot's policy $\pi^r\colon \mathcal{S} \times \mathcal{A}^r \to [0,1]$, where $\mathcal{A}^r$ denotes the set of possible robot actions. $\pi^r(\mathbf{s}, a^r)$ specifies the probability of choosing action a^r in state $\mathbf{s}$.

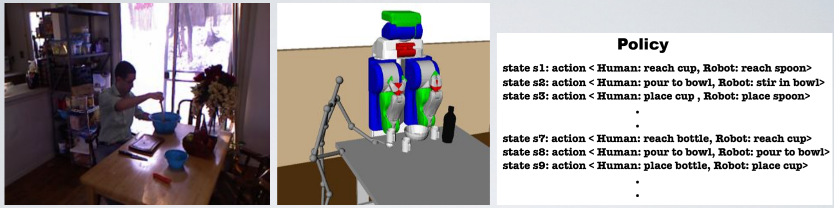


Fig. 3: **Collaborative planning by the robot.** In order to collaborate with the human on a *recipe following* task, the robot learns the activity model from RGB-D videos of human preparing a recipe (left), represents the environment via affordances and uses our planning algorithm (middle) to generate a policy for jointly performing the activity with the human (right).

State-Action Representation

- represent the environment in terms of the object affordances; one affordance labels of the objects in the scene
  - bowl is stirrable,
  - spoon is the stirrer, …
- e.g. stir action corresponds to the temporal motion trajectory of the spoon. On performing the stir action, the spoon becomes placeable, thus changing the state of the environment.

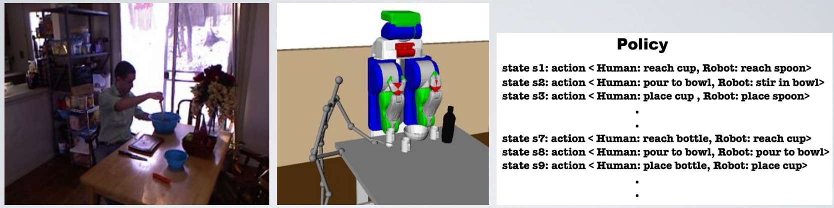


Fig. 3: **Collaborative planning by the robot.** In order to collaborate with the human on a *recipe following* task, the robot learns the activity model from RGB-D videos of human preparing a recipe (left), represents the environment via affordances and uses our planning algorithm (middle) to generate a policy for jointly performing the activity with the human (right).

Task Model: State Transitions and Rewards

- on performing a valid action, the affordance of the object changes to another fixed affordance
- reward function allows us specify valid actions at any given affordance state, where all valid actions receive a fixed positive reward and non-valid actions will incur a negative cost

Algorithm 1 RUN-EPISODE (Q, π^h, π^r)

INPUT: State space S, Action space A

```
Initialize environment to start state;   R ← 0;   i ← 0;
loop
   if goal state then
      return Q-functions, R
   end if
   Sample a_h from π^h and a_r from π^r
   Take action (a_h, a_r) and observe r, s'
   Update Q-functions as in Eq. 2
   R ← R + γ^i * r
   i ← i + 1
end loop
```

Learning Robot Policy

- *Q-learning* algorithm to learn the local value functions $q_t^h(s,a)$ and $q_t^r(s,a)$ for the human and the robot respectively.
- take joint actions, advance when both actions completed (there is a "do nothing" action)

MODELS OF HUMAN BEHAVIOUR

Habit-following human.

- modelled trained from human alone
- assume that the human follows close to what he has done in the training videos.
- fixed strategy behaviour, where the human has a preferred way of performing activities and follows the same approach even when working with a robot.
- Let D be the set of activity videos and let c(s,a) be the number of times the human performed action a when in state s in D.

$$\pi_d^h(\mathbf{s}, a_i) = \begin{cases} c(\mathbf{s}, a_i)/\sum_a c(\mathbf{s}, a) & \text{if } \mathbf{s} \in D \\ 1/n & \text{if } \mathbf{s} \notin D \\ 0 & \text{otherwise} \end{cases}$$

MODELS OF HUMAN BEHAVIOUR

ϵ-optimal human.

- assume that the human takes the best action according to the value function most of the time, but makes a random choice ϵ fraction of the time.
- human chooses a response that is mutually beneficial most of the time, according to the value function learnt so far.

$$\pi^h_\epsilon(\mathbf{s}, a_i) = \begin{cases} (1-\epsilon) + (\epsilon/n) & \text{if } a_i = argmax_a(q^h(\mathbf{s}, a)) \\ \epsilon/n & \text{otherwise} \end{cases}$$

MODELS OF HUMAN BEHAVIOUR

Adaptive human.

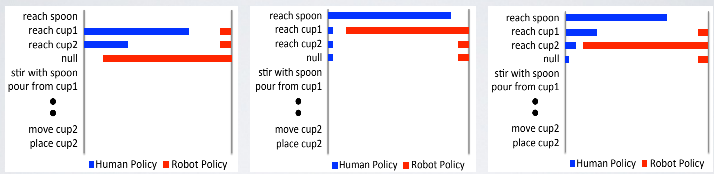
- Humans usually adapt to other agents while trying to maintain their preferences or habits.
- follow habits when possible in familiar situations, but when faced with new situations while working with the robot, adapt and try to perform the action that is beneficial to both for completing the activity.

$$\pi^h_a(\mathbf{s}, a) = \eta * \pi^h_d(\mathbf{s}, a) + (1-\eta) * \pi^h_\epsilon(\mathbf{s}, a)$$

$$\pi^h_\epsilon(\mathbf{s}, a_i) = \begin{cases} (1-\epsilon) + (\epsilon/n) & \text{if } a_i = argmax_a(q^h(\mathbf{s}, a)) \\ \epsilon/n & \text{otherwise} \end{cases}$$

$$\pi^h_d(\mathbf{s}, a_i) = \begin{cases} c(\mathbf{s}, a_i)/\sum_a c(\mathbf{s}, a) & \text{if } \mathbf{s} \in D \\ 1/n & \text{if } \mathbf{s} \notin D \\ 0 & \text{otherwise} \end{cases}$$

RESULTS



(a) Habit-following Human. (b) ϵ-optimal Human. (c) Adaptive Human.

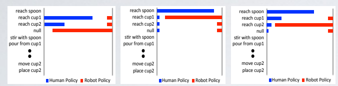
Fig. 4: **Illustration of policies learnt with different human behaviors:** Each figure shows the learnt probability distributions of the various possible actions at the start state of the *following recipe* activity. Blue and red bars represent the probability of choosing the corresponding actions by the human and robot respectively.

$$\pi^h_a(\mathbf{s}, a) = \eta * \pi^h_d(\mathbf{s}, a) + (1-\eta) * \pi^h_\epsilon(\mathbf{s}, a)$$

$$\pi^h_\epsilon(\mathbf{s}, a_i) = \begin{cases} (1-\epsilon) + (\epsilon/n) & \text{if } a_i = argmax_a(q^h(\mathbf{s}, a)) \\ \epsilon/n & \text{otherwise} \end{cases}$$

$$\pi^h_d(\mathbf{s}, a_i) = \begin{cases} c(\mathbf{s}, a_i)/\sum_a c(\mathbf{s}, a) & \text{if } \mathbf{s} \in D \\ 1/n & \text{if } \mathbf{s} \notin D \\ 0 & \text{otherwise} \end{cases}$$

RESULTS



(a) Habit-following Human. (b) ϵ-optimal Human. (c) Adaptive Human.

Fig. 4: **Illustration of policies learnt with different human behaviors:** Each figure shows the learnt probability distributions of the various possible actions at the start state of the *following recipe* activity. Blue and red bars represent the probability of choosing the corresponding actions by the human and robot respectively.

Table 1: **Collaborative Planning Evaluation.** Metrics computed for the collaborative plans generated on our RGB-D dataset.

Model	% time saving					% conflicts
	Recipe	Setting	Cleaning	Loading	Overall	
Human Expert Plans	36.8	53.1	16.4	42.4	37.2	0
Chance	3.3	10.5	-33.1	23.7	1.1	3.7
Mental-model MDP[33]	-2.6	30.4	-5.1	32.3	13.8	6.4
Our Model – ϵ-optimal human	27.5	45.6	18.3	30.8	31.2	13.5
Our Model – habit following human	28.4	48.1	18.6	41.4	33.4	11.9
Our Model – adaptive human	**32.8**	**48.5**	**22.9**	**41.9**	**36.5**	13.7

RESULTS

Table 1: **Collab** What's this? Metrics computed for the collaborative plans gener...

Model	Time saving					% conflicts
	Recipe	Setting	Cleaning	Loading	Overall	
Human Expert Plans	36.8	53.1	16.4	42.4	37.2	0
Chance	3.3	10.5	-33.1	23.7	1.1	3.7
Mental-model MDP[33]	-2.6	30.4	-5.1	32.3	13.8	6.4
Our Model – ϵ-optimal human	27.5	45.6	18.3	30.8	31.2	13.5
Our Model – habit following human	28.4	48.1	18.6	41.4	33.4	11.9
Our Model – adaptive human	**32.8**	**48.5**	**22.9**	**41.9**	**36.5**	13.7

Nikolaidis, S. & Shah, J., 2013. Human-robot cross-training: Computational formulation, modeling and evaluation of a human team training strategy. In ACM/IEEE International Conference on Human-Robot Interaction. pp. 33–40.

ANOTHER MDP-BASED MODEL

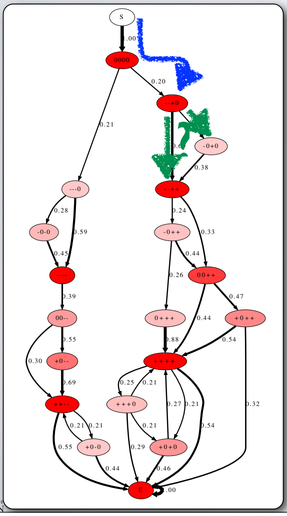
What's the difference?

Nikolaidis, S. & Shah, J., 2013. Human-robot cross-training: Computational formulation, modeling and evaluation of a human team training strategy. In ACM/IEEE International Conference on Human-Robot Interaction. pp. 33–40.

We describe how the robot teaming mode... tationally encoded as a Markov Decision Pr... decision process is a tuple $\{S, A, T, R\}$, where:

- S is a finite set of states of the world; it models the set of world environment configurations.
- A is a finite set of actions; this is the set of actions the robot can execute.
- $T : S \times A \longrightarrow \Pi(S)$ is the state-transition function, giving for each world state and action, a probability distribution over world states; the state transition function models the variability in human action. For a given robot action a, the human's next choice of action yields a stochastic transition from state s to a state s'. We write the probability of this transition as $T(s, a, s')$. In this formulation, human behavior is the cause of randomness in our model, although this can be extended to include stochasticity from the environment or the robot actions, as well.
- $R : S \times A \longrightarrow \mathbb{R}$ is the reward function, giving the expected immediate reward gained by taking each action in each state. We write $R(s, a)$ for the expected reward for taking action a in state s.

ANOTHER MDP-BASED MODEL



We describe how the robot teaming model can be computationally encoded as a Markov Decision Process. A Markov decision process is a tuple $\{S, A, T, R\}$, where:

- S is a finite set of states of the world; it models the set of world environment configurations.
- A is a finite set of actions; this is the set of actions the robot can execute.
- $T : S \times A \longrightarrow \Pi(S)$ is the state-transition function, giving for each world state and action, a probability distribution over world states; the state transition function models the variability in human action. For a given robot action a, the human's next choice of action yields a stochastic transition from state s to a state s'. We write the probability of this transition as $T(s, a, s')$. In this formulation, human behavior is the cause of randomness in our model, although this can be extended to include stochasticity from the environment or the robot actions, as well.
- $R : S \times A \longrightarrow \mathbb{R}$ is the reward function, giving the expected immediate reward gained by taking each action in each state. We write $R(s, a)$ for the expected reward for taking action a in state s.

ANOTHER MDP-BASED MODEL

robot has policy and chooses action, learn what human does

Algorithm : Human-Robot Cross-training
1. Initialize $R(s, a)$ and $T(s, a, s')$ from prior knowledge
2. Calculate initial policy π
3. **while**(number of iterations < *MAX*)
4. Call Forward-phase(π)
5. Update $T(s, a, s')$ from observed sequence $s_1, a_1, s_2, \dots, s_{M-1}, a_{M-1}, s_M$
6. Call Rotation-phase()
7. Update $R(s_i, a_i)$ for observed sequence $s_1, a_1, s_2, a_2, \dots, s_N, a_N$
8. Calculate new policy π
9. **end while**

learn reward based on what human would do!

We describe how the robot teaming model can be computationally encoded as a Markov Decision Process. A Markov decision process is a tuple $\{S, A, T, R\}$, where:

- S is a finite set of states of the world; it models the set of world environment configurations.
- A is a finite set of actions; this is the set of actions the robot can execute.
- $T : S \times A \longrightarrow \Pi(S)$ is the state-transition function, giving for each world state and action, a probability distribution over world states; the state transition function models the variability in human action. For a given robot action a, the human's next choice of action yields a stochastic transition from state s to a state s'. We write the probability of this transition as $T(s, a, s')$. In this formulation, human behavior is the cause of randomness in our model, although this can be extended to include stochasticity from the environment or the robot actions, as well.
- $R : S \times A \longrightarrow \mathbb{R}$ is the reward function, giving the expected immediate reward gained by taking each action in each state. We write $R(s, a)$ for the expected reward for taking action a in state s.

ANOTHER MDP-BASED MODEL

Algorithm : Human-Robot Cross-training
1. Initialize *R*(*s*, *a*) and *T*(*s*, *a*, *s'*) from prior knowledge
2. Calculate initial policy π
3. **while**(number of iterations < *MAX*)
4. Call Forward-phase(π)
5. Update *T*(*s*, *a*, *s'*) from observed sequence $s_1, a_1, s_2, \ldots, s_{M-1}, a_{M-1}, s_M$
6. Call Rotation-phase()
7. Update R(s_i, a_i) for observed sequence $s_1, a_1, s_2, a_2, \ldots, s_N, a_N$
8. Calculate new policy π
9. **end while**

Function: Forward-phase(policy π)
1. Set *current_state* = START_STATE
2. **while**(*current_state* != FINAL_STATE)
3. Execute robot action *a* according to current policy π
4. Observe human action
5. Set *next_state* to the state resulting from *current_state*, robot and human action
6. Record *current_state*, *a*, *next_state*
7. *current_state* = *next_state*
8. **end while**

Function: Rotation-phase()
1. Set *current_state* = START_STATE
2. **while** (*current_state* != FINAL_STATE)
3. Set action *a* to observed human action
4. Sample robot action from *T*(*current_state*, *a*, *next_state*)
5. Record current_state, *a*
6. *current_state* = *next_state*
7. **end while**

RESULTS

1) Cross-training session (Group A): The participant iteratively switches positions with the virtual robot, placing the screws at the forward phase and drilling at the rotation phase.
2) Reinforcement learning with human reward assignment session (Group B): This is the standard reinforcement learning approach, where the participant places screws and the robot drills at all iterations, with the participant assigning a positive, zero, or negative reward after each robot action [10].

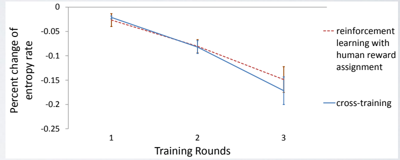


Fig. 7. Human-Robot Mental Model Convergence. The graph shows the percent decrease of entropy rate over training rounds.

Algorithm : Human-Robot Cross-training
1. Initialize *R*(*s*, *a*) and *T*(*s*, *a*, *s'*) from prior knowledge
2. Calculate initial policy π
3. **while**(number of iterations < *MAX*)
4. Call Forward-phase(π)
5. Update *T*(*s*, *a*, *s'*) from observed sequence $s_1, a_1, s_2, \ldots, s_{M-1}, a_{M-1}, s_M$
6. Call Rotation-phase()
7. Update R(s_i, a_i) for observed sequence $s_1, a_1, s_2, a_2, \ldots, s_N, a_N$
8. Calculate new policy π
9. **end while**

$$H(X) = -\sum_{s \in S} \mu(s) \sum_{s' \in S} T(s, \pi(s), s') log\left[T(s, \pi(s), s')\right]$$

Human-Robot Cross-Training: Computational Formulation, Modeling and Evaluation of a Human Team Training Strategy

Stefanos Nikolaidis
Department of Aeronautics and Astronautics
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, MA 02139
Email: snikol@mit.edu

Julie Shah
Department of Aeronautics and Astronautics
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, MA 02139
Email: julie_a_shah@csail.mit.edu

5-3 Sisbot, E.A. & Alami, R., 2012. A human-aware manipulation planner. IEEE Transactions on Robotics, 28(5), pp.1045–1057.	Paolo Russo	Ahmad Irjoob

4-2 (*moved here due to unavailability of presenter next week*) Lang, C. et al., 2009. Feedback interpretation based on facial expressions in human-robot interaction. In RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication. Toyama, Japan: IEEE, pp. 189–194.	Wilson Villa	Harold Agudelo