Artificial Cognitive Systems

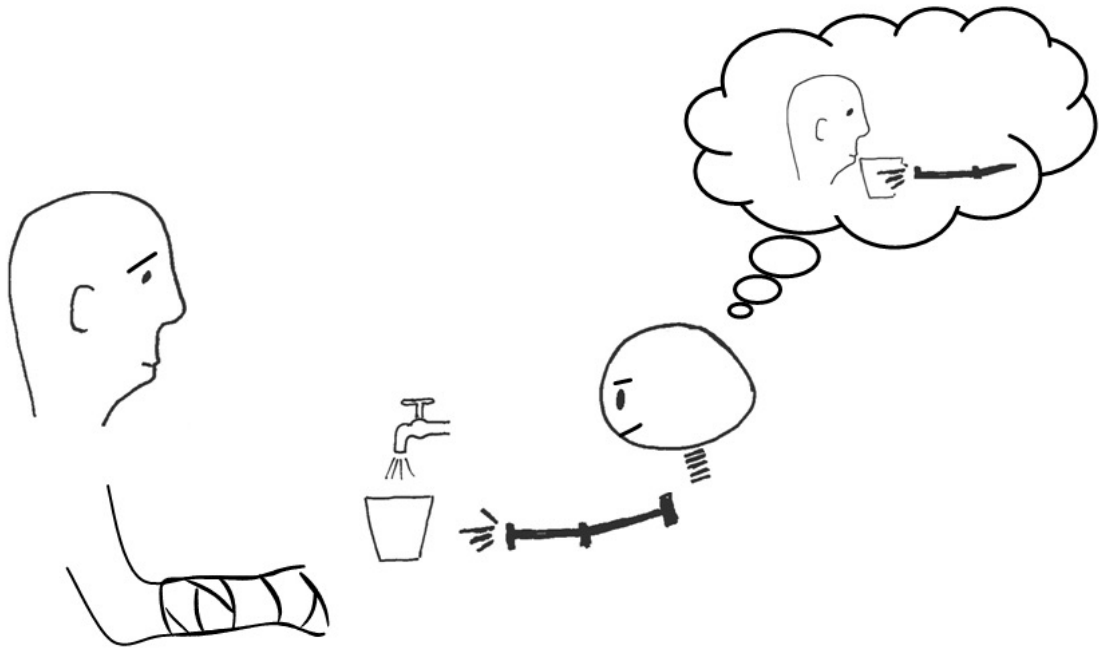
Module 9: Social Cognition

Lecture 2: Reading intentions and theory of mind; instrumental helping; collaboration

David Vernon
Carnegie Mellon University Africa

www.vernon.eu

Social Cognition



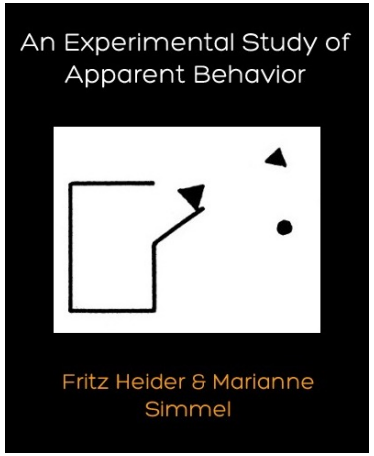
Inferring Intentions & Theory of Mind
Instrumental Helping
Collaboration

Inferring Intentions and Theory of Mind

- Prospection is the essence of cognition
- Social interaction is complex because …
  - **A cognitive agent's act must anticipate the actions of an agent that itself is already anticipating what it is going to do**
- That is, an agent must anticipate the intentions of other agents
  - Predict what they will do
  - Possibly, why they want to do it

Inferring Intentions and Theory of Mind

- Theory of Mind
  - The ability to infer what someone else is thinking and wants to do
- Young children differentiate between the behaviour of inanimate and animate objects
  - Attribute mental states to the animate objects
- Attribute agency (animacy) to inanimate objects that exhibit biological motion
  - Intentions, emotions, motives, personality traits



Fritz Heider and Marianne Simmel, 1944

https://www.youtube.com/watch?v=n9TWwG4SFWQ
and
http://vimeo.com/48908599

Inferring Intentions and Theory of Mind

Humans infer different types of intention

- Interpreting movements (lower level intentions) … **what** is the desired state
- Interpreting actions (higher level) … **why** is it desired (underlying motive)
- Grasp a cup *vs.* he is thirsty

Inferring Intentions and Theory of Mind

- How do human infer the intentions of others from their actions?

- **Internal simulation**

  - A mechanism that predicts the consequences of the agent's own actions based on its own intentions by internal simulation using forward models

    - Input: overt or covert motor commands
    - Output: the likely sensory consequences of carrying out those commands

  - When observing another agent's actions, the same mechanism can be used

    - Associate **observed** movements with likely, i.e. intended, sensory consequences

Inferring Intentions and Theory of Mind

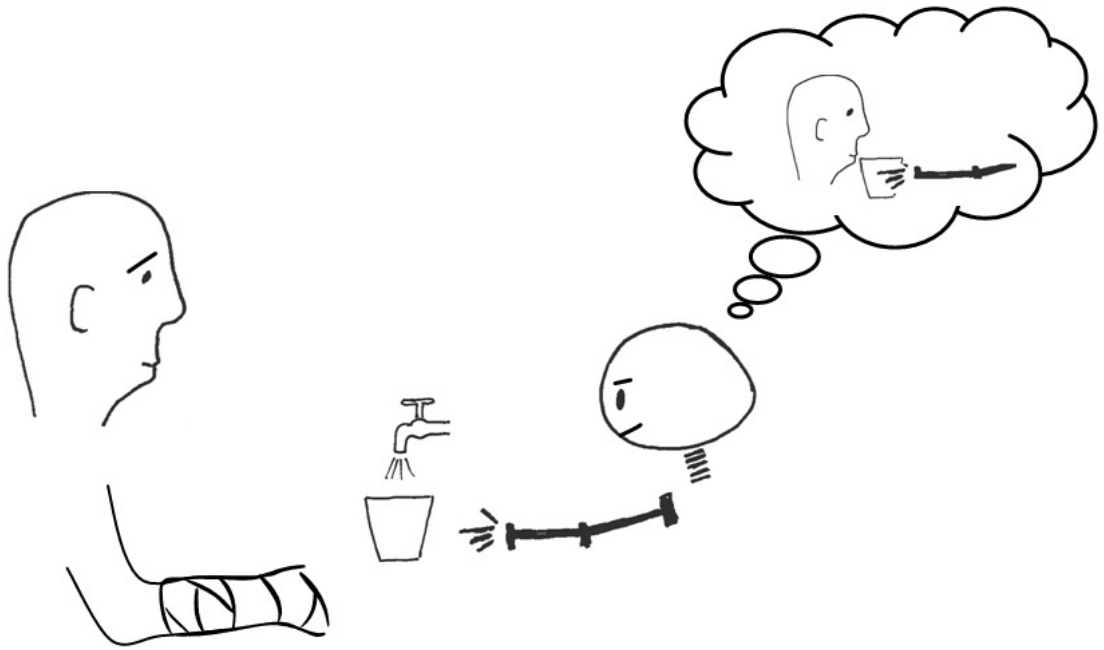
Internal simulation

- Ideo-motor models
  - Focus on goals and intentions
  - Use a joint representation that embraces both perception and action
- When an agent just sees another agent's action
  - The agent's own actions activated
  - And so too are the consequences of those actions
  - Hence the intention of the actions can be inferred
- With a suitably-sophisticated joint representation and internal simulation mechanism, both low-level movement intentions and high-level action intentions can be accommodated

Inferring Intentions and Theory of Mind

Internal simulation

- HAMMER can be used to give robots the ability to read intentions
  - By internal simulation to form a theory of mind
  - Exploiting the multiple pairs of inverse and forward models as a correlate of the mirror neuron system and a realization of ideo-motor theory



Inferring Intentions & Theory of Mind
Instrumental Helping
Collaboration

Instrumental Helping

Learning to Help and Be Helped

It takes several years for human infants to develop the requisite abilities

Newborns
Newborns gaze longer when a person looks directly at them [256].
Newborns are attracted to people (i.e. face and voice) [257].
Newborns prefer biological motion [258].
Newborns preferentially orient toward faces [259, 260].
Newborns prefer human voices to other sounds [261].
Early Development
$2\frac{1}{2}$ months: infants can discriminate a familiar adult's expressions if presented with multimodal expressions [262].
3 months: infants engage mutual gaze with adults, i.e. both agents attend to each other's eyes simultaneously [263].
3–4 months: infants have the ability to discriminate among a few photographed, static facial expressions [264].
4 months: infants presented with multimodal expressions can discriminate some adult's expressions [265, 266].
5 months: infants discriminate auditory-only displays of affect [266].
6 months: infants can perceive approximate direction of attention of others (i.e. to the left or to the right) [267].
10–12 months: infants show the first strong evidence of understanding the feelings of others.
9 months: infants can accurately detect the direction of the adult's gaze [263].
12 months: infants look at the object fixated by the adults [268].
12 months: infants consider eye rather than head direction [269].
12 months: Children start to understand pointing as an object-directed action [270].
12 months: Children anticipate with gaze the goal of a feeding action [271].
Later Development
18 months: children start to follow an adult's gaze outside their own field of view [263].
18 months: children perceive from emotions that a person wants something [272].
18 months: infants can infer what another person is trying to achieve (even if the attempt is unsuccessful) [273, 274].
18 months: infants altruistically (*instrumentally*) help adults when they are having problems achieving a goal [275].

Instrumental Helping

Learning to Help and Be Helped

- 14 to 18 months: **instrumental helping behaviour**
- 2 years: start to solve **simple cooperation** tasks with adults
- 2-3 years: ability to **cooperate** with peers
- 3 years: complex **collaboration**
  - Sharing of intentions
  - Joint coordination of complementary actions
- 3 ½ years: Roles in the task can be reversed; & can teach new partners

Instrumental Helping

Young children (one year old) are naturally altruistic

- Innate propensity to help others even when no reward is offered
- Comforting another person who appears to be distressed
- Helping someone achieve something that they can't do on their own

  (e.g. picking up something they've dropped when their hands are full)

Instrumental Helping

Instrumental helping

- Assist the other individual achieve their goals
- Even in the absence of any benefit from doing so
- Sometimes because there is no benefit in doing so

Instrumental Helping

The Prospective Nature of Helping

- **Instrumental helping** requires one agent to understand the goal of another agent
  - Inferring its intention
  - Recognizing that it can't achieve it without assistance
  - Acting to provide the necessary help

    (e.g., picking up something for a person whose hands are full).

Instrumental Helping

- Two components

  1. Cognitive

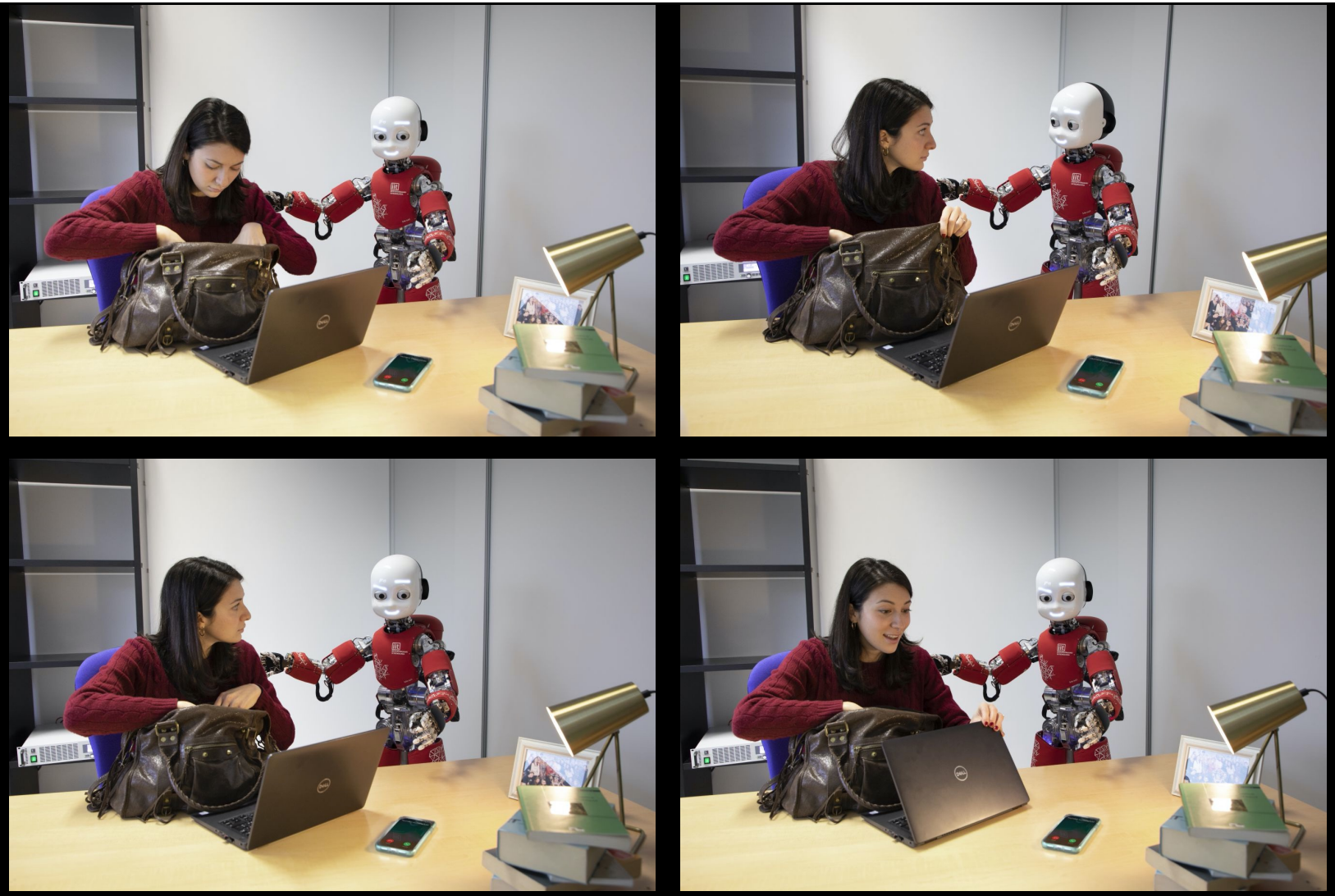
     - Recognizing what the other agent's goal is
     - Recognizing the reason they can't achieve it on their own

  2. Emotional

     - Desire to see the second agent achieve the goal
     - Desire to see the second agent exhibit pleasure at achieving it

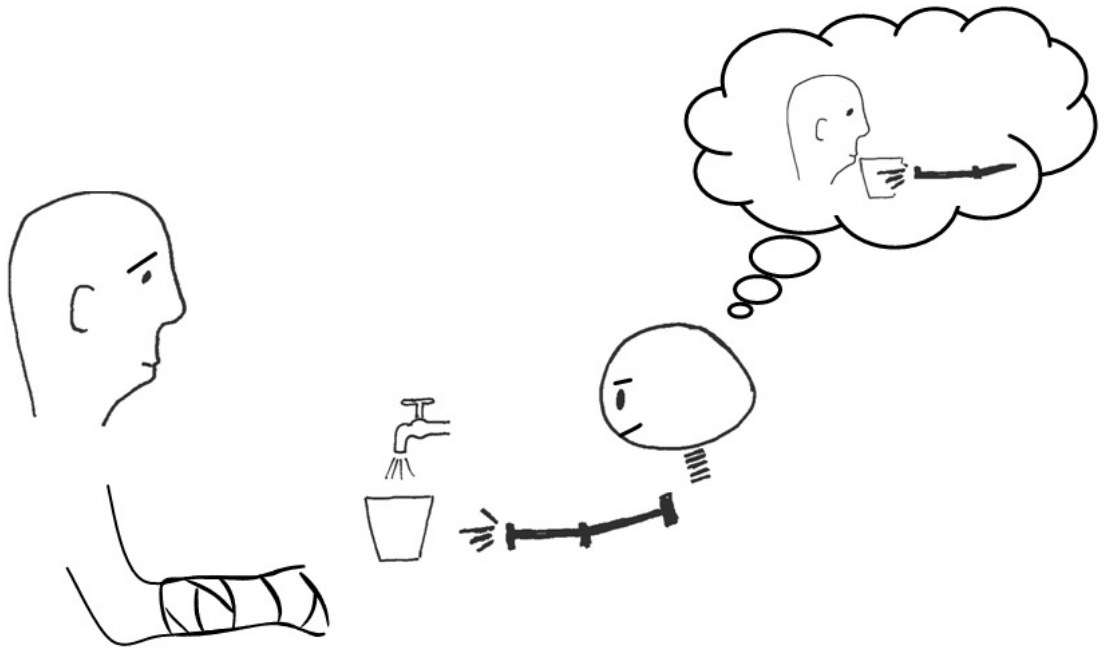
- Prosocial behaviour (in contrast to anti-social behaviour)

  – Directed at benefitting another person



"This sequence of pictures depicts a situation in which the iCub humanoid robot (www.icub.org) is interacting with a human, reading her intention to get her phone from her bag, and alerting her to the fact that it is on the desk, hidden from her by the laptop.
Note that this sequence has been staged to illustrate the desired capabilities of a cognitive robot and has not yet been implemented."

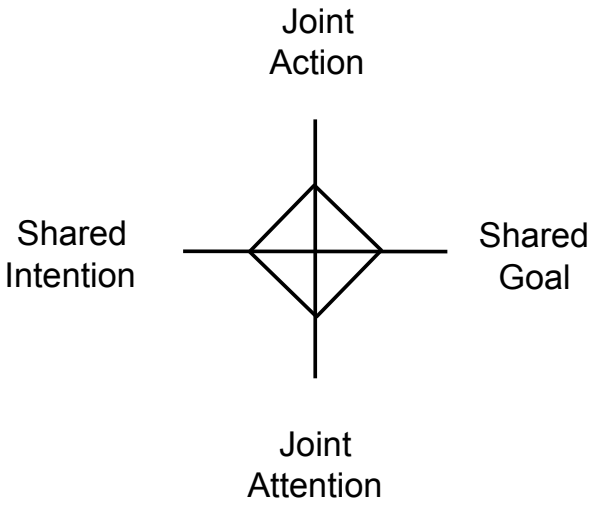
Sandini, G., A. Sciutti, and D. Vernon (in press). Cognitive Robotics. In M. Ang, O. Khatib, and B. Siciliano (Eds.), Encyclopedia of Robotics. Springer.
Images courtesy of Istituto Italiano di Tecnologia

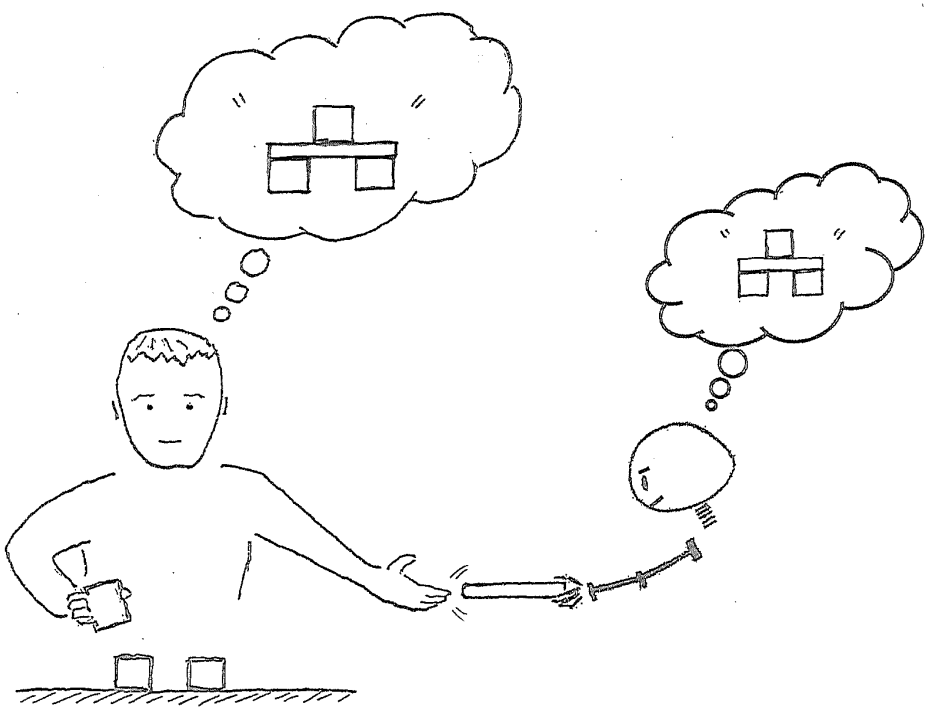


Inferring Intentions & Theory of Mind
Instrumental Helping
Collaboration

Collaboration

- More complicated than instrumental helping
- **Joint cooperative action**
  - Joint action
  - Shared cooperative activity
- Agents that engage in joint action
  - **Share** the **same goal, intend** to act together, **coordinate** their actions to achieve their shared goal through **joint attention**


Joint
Action
Shared
Intention
Shared
Goal
Joint
Attention



Collaboration

- To take part in collaborative activities requires
  - An ability to read intentions
  - An ability to infer goals
  - A unique motivation to share psychological states with other agents
- Shared intentionality:

"Collaborative actions in which participants have a shared goal (shared commitment) and coordinated action roles for pursuing that shared goal."

Tomasello et al. 2005

Collaboration

- The goals and intentions of each agent is a mix of their own goals & intentions and the goals and intentions of the other agent
  - The intention is a shared intention
  - The associated actions are joint actions
- This differentiates collaboration from instrumental helping

Collaboration

- Furthermore, each agent understands both roles of the interaction and so can help the other agent if required

- Critically

  - agents not only choose their own action plan

  - but also, in its own motor system to enable coordination,
    represent the other agent's action plan
    in the sense of who is doing what and when


KUKA
KUKA

Collaboration

Joint Action

There are at least six degrees of freedom in joint action

1. The number of participants involved
2. The nature of the relationship between the participants (e.g. peer-to-peer or hierarchical)
3. Whether or not the roles are interchangeable
4. Whether the interaction is physical or virtual
5. Whether or not the participants' association is temporary or more long-lasting
6. Whether or not the interaction is regulated by organizational or cultural norms

Collaboration

Joint Action

- Joint action, or shared cooperative activity, has three essential characteristics

    1. Mutual responsiveness
    2. Commitment to joint activity
    3. Commitment to mutual support

- Each agent must be mutually responsive to the intentions and actions of the other
- Each must know that the other is trying to be similarly responsive
- Consequently, each agent behaves in a way that is guided partially by the behaviour of the other agent
- Not the same as instrumental helping

Collaboration

Joint Action

- Each agent must also be committed to the activity in which they are engaged
- Both agents have the same intention but they need not have the same reason for engaging in the activity
- This is a subtle point:
  - The outcome of the collaboration is the same for both agents
  - But the reason for adopting the goal of achieving that outcome need not be the same

Collaboration

Joint Action

If a cognitive robot and a disabled person collaborate to do the laundry

- The outcome — the goal — may be a wardrobe full of clean clothes
- The reason the person has the goal is to have a fresh shirt to wear in the morning
- The reason the robot has the goal may just be to keep the house clean and uncluttered

Collaboration

Joint Action

- Each agent must be committed to supporting the efforts of the other agent to play their role in the joint activity
  - Each agent will provide any help the other agent requires
  - Each agent treats this collaborative mutual support as a priority activity
- Shared intentions are essential for joint action
  - The intentions of each agent must mesh together in a mutually-supportive manner

Collaboration

Shared Intentions

- We-intention
- Collective intention
- Joint intention

- Not simply a collection of individual intentions
  - Agents with a shared intention represent the overall shared goal between them
  - Each agent only has its own partial **sub-plans**
  - An individual agent with a shared intention does not need to know the other agent's partial **plan**
  - However, they do need to share the overall **goal**

Collaboration

Shared Intentions

- Realization of a shared intention and the execution of a joint action
  - Each agent must represent its own actions and their predicted consequences
  - AND the goals, intentions, actions and predicted consequences of the other agent
  - AND the effect that their actions have on the other agent
  - AND have at least a partial representation of how actions combine to achieve the overall goal
  - AND have an ability to predict the effects of their joint actions ← so that it can monitor progress towards the overall goal and adjust its actions to help the other agent
  - AND have shared perceptual representation (joint attention)

Collaboration

Joint Attention

- Joint attention involves much more than two agents looking at the same thing
- Joint attention: the relationship between intentionality and attention
- Joint attention is

"(1) a coordinated and collaborative coupling between intentional agents where (2) the goal of each agent is to attend to the same aspect of the environment"

Kaplan and Hafner 2006

Collaboration

Joint Attention

- Joint attention requires shared intentionality
- The participating agents must be engaged in collaborative intentional action
- Each agent must monitor, understand, and direct the attentional behaviour of the other agent
- Both agents must be aware that this is going on

Collaboration

Joint Attention

- Joint attention is an on-going mutual activity that is carried on throughout the collaborative process to monitor and direct the attention of the other agent

- Joint attention is, itself, a joint activity

Collaboration

Joint Attention

Four skills required by an agent

1. Detect and track the attentional behaviour of the other agent

Collaboration

Joint Attention

Four skills required by an agent

2. Influence the attentional behaviour of the other agent

    - possibly by using gestures such as pointing or by use of appropriate words

Collaboration

Joint Attention

Four skills required by an agent

3. Engage in social coordination to manage the interaction

    - Taking turns or swapping roles

Collaboration

Joint Attention

Four skills required by an agent

4. Understand that the other agent has intentions

   - Could be different from its own provided the goal is the same
   - Intentional understanding: able to interpret and predict the behaviour of the other agent in terms of the actions required to reach the shared goal

Recommended Reading

Vernon, D. Artificial Cognitive Systems – A Primer, MIT Press, 2014; Chapter 9.

D. Vernon, S. Thill, and T. Ziemke, "The Role of Intention in Cognitive Robotics", in Toward Robotic Socially Believable Behaving Systems- Volume I, A. Esposito and L. C. Jain (Eds.), Intelligent Systems Reference Library, Vol. 105, pp. 15-27, Springer, 2016.