

JOINT ATTENTION

Kaplan and Hafner (2006)

Florian Niefind

Coli, Universität des Saarlandes

SS 2010

1. Outline

2. Joint attention - an informal approximation

3. Motivation of the paper

4. Formalization of the problem

5. Developmental timeline and artificial models

(a) Attention detection

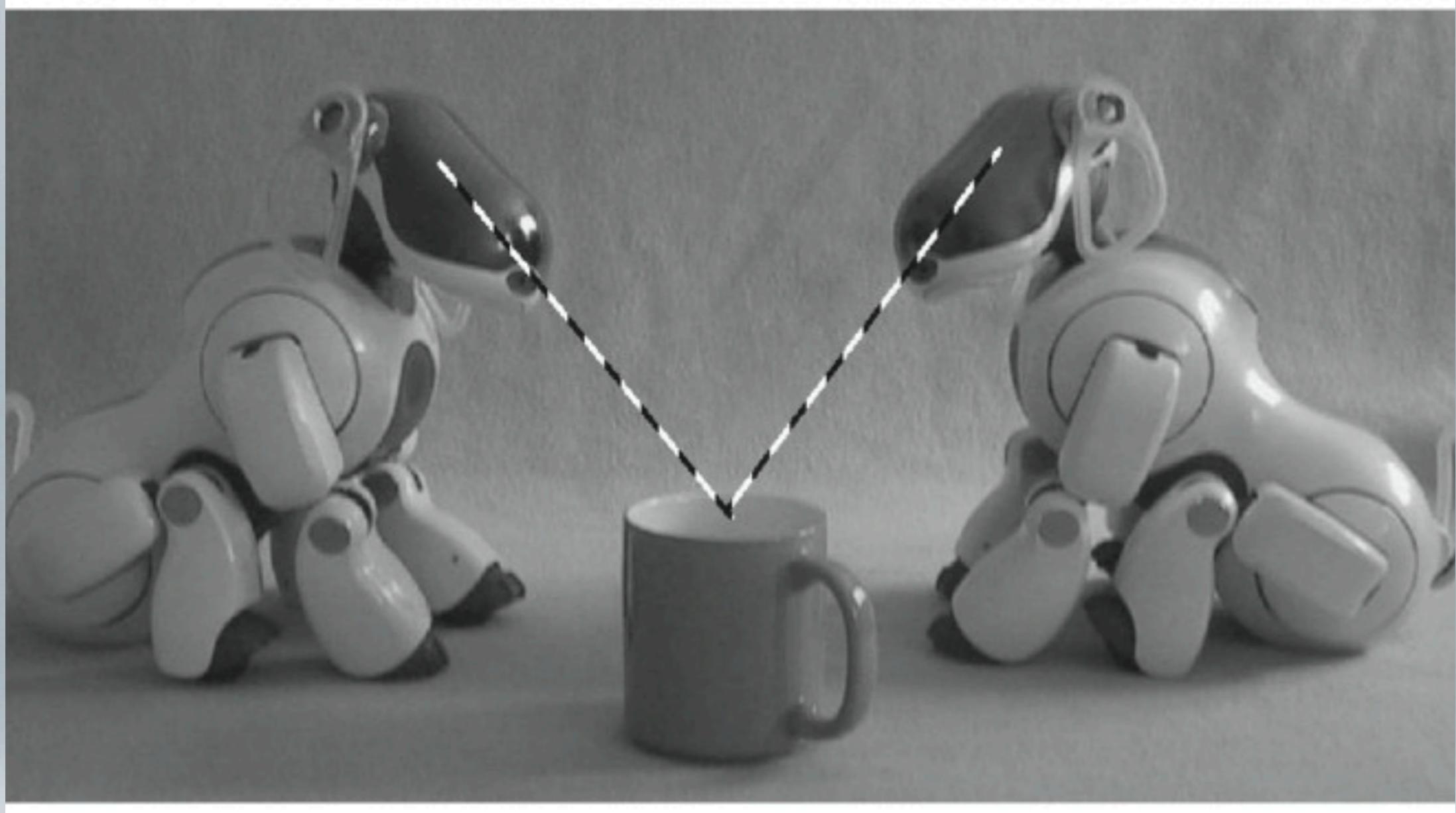
(b) Attention manipulation

(c) Social coordination

(d) Intentional Understanding

(e) Integration of the skills

6. Conclusions



JOINT ATTENTION -INFORMAL

JOINT ATTENTION - INFORMAL

- joint attention is an important issue in many current fields of research
 - developmental psychology
 - robotics
 - (interactional) linguistics
- however there is no clear formal definition that is used in all these fields
- we are going to see that from four examples from Kaplan and Hafner (2006)

SIMULTANEOUS LOOKING

- Two robots are sitting in a room. Suddenly one of their toys makes a squeaking noise. They both turn and look at it.
- passive attention triggered by a salient event (can also be triggered by e.g. pop-out effects)

COINCIDENTAL SIMULTANEOUS LOOKING

- The robots are looking for a toy to play with. At the same moment, they both see a pink ball on the floor. They pay attention to it without noticing each other. Each other's behavior is not monitored.
- active attention, but no attention detection, no intentional joint attention

GAZE FOLLOWING

- One robot is looking at a new toy. The other less experienced robot follows his gaze since it has learned (for instance with a reinforcement learning algorithm) that by doing that, it will often see something salient.
- active attention, but no attention detection, no intentional joint attention

COORDINATED GAZE ON AN OBJECT

- Both robots are looking at a toy bunny, and are also informed that the other one is looking, too. One robot is attending to the bunny in order to play with it, the other one is purely attracted by its color.
- active attention, but attention is not joint, as they are not attending to the same aspect of the object

MOTIVATION OF THE PAPER

- as we can see there are a lot of criteria for joint attention
- this leads to different definitions of the term in different fields of research
- especially in social robotics it is very ill-defined, what problem needs to be solved
 - concentration on partial and isolated elements of joint attention
 - often just simultaneous looking or simple coordinated behavior
 - deeper cognitive issues are not adressed

OUTLINE

- present a clear formal account of the joint attention concept
- identify the different skills underlying it's development
- review experimental observations from developmental psychology regarding those skills
- review corresponding relevant models of robotics
- show directions for further investigations

FORMALIZATION OF THE PROBLEM

- the paper follows the view of Tomasello (1995) that attention can only be understood through it's relations with intentional actions
- from this starting point we can adress the following issues:
 - What is attention?
 - How is attention used to perform intentional actions?
 - How can these actions be observed and interpreted by an external agent?
 - How can interaction between intentional agents be modeled?
 - What characterizes **joint** attention?

ATTENTION

- temporally extended process
- agent concentrates on certain features of the environment
- thereby excluding other features (relatively)

TWO KINDS OF ATTENTION

- there are two groups of factors that make children attend to something
 - occurrence of salient events: bottom-up or passive attention
 - active direction of attention by the child: top-down or active attention
- experimental evidence suggests two different neurological systems, the bottom-up system can serve as an interrupt device for the top-down system

GOALS AND INTENTIONS

- activities that require focus on aspects of the environment are **goal-directed** or **intentional** processes
- goal: situation that the agent tries to achieve
- intention: plan of action to realize the goal
 - includes the goal as well as the means to achieve it

ATTENTION AS INTENTIONALLY DIRECTED PERCEPTION

- to realize the goal the agent focusses on relevant perceptual features
- the agent evaluates the efficiency of the action plan towards the goal
- thus attention is intentionally-directed perception
- it provides means to assess the effectiveness (success or failure) and efficiency of an attempt to reach the goal
- self-regulating systems can be described as performing intentional actions by this definition

OBSERVATION OF INTENTIONAL ACTIONS

- goals and action plans are not observable directly!
- intentional actions are associated with observable effects (cf. table)
 - movements
 - attentional behavior
 - verbal responses
 - joy if goal is reached
 - aim of behavior can eventually be observed in the environment

Internal processes	Observable behavior
Visual Attention	Gaze direction
Auditory Attention	Head orientation
Action plan	Observed movements
Goal	Actual transformation of the environment
Progress, success, failure	Emotional expressions, surprise

INTERPRETATIONS OF OBSERVATIONS

- these observations need to be interpreted
- this interpretation process is far from being understood
- it seems to be based on the perceptual cues we just saw
- there is also most probably a matching process with ones own actions
- neuronal evidence might come from mirror neurons

COUPLING BETWEEN INTENTIONALLY DRIVEN PROCESSES

- the **interaction** of intentional agents makes everything way more complex
 - the action plan of one agent can take the observed behavior of the other agent into account
 - his behavior can influence the others behavior
 - observations can be misinterpreted
 - it can be a goal to **change** the other agents behavior
 - it can be a goal to **coordinate** activities

JOINT ATTENTION

- joint attention is often defined as two agents looking at the same thing
- if we reconsider the examples from the beginning we see that this is not necessarily true

Case	Active / Passive	Attention detection	Unilateral / Bilateral
Case 1: Simultaneous looking triggered by a salient event or a “pop-out” effect	Passive	No	–
Case 2: Coincidental simultaneous looking	Active	No	–
Case 3: Gaze following	Active	Yes	Unilateral
Case 4: Coordinated gaze on same object	Active	Yes	Bilateral

DEFINING JOINT ATTENTION

- a coordinated and collaborative coupling between intentional agents
- the **goal** of each agent is to attend to the same aspect of the environment

Tomasello (2004):

Joint attention is an active bilateral process which involves attention alternation, but it can only be fully understood if we assume that it is realized by agents performing intentional actions. To achieve joint attention, agents must monitor, influence and coordinate their behavior in order to engage in a collaborative intentional action. They must reach what Tomasello calls a form of shared intentionality

4 PREREQUISITES FOR JOINT ATTENTION

- **attention detection**
 - track attentional behavior of other agents
- **attention manipulation**
 - manipulate and influence other agents attentional behavior
- **social coordination**
 - engage in coordinated activities with other agents
- **intentional understanding**
 - view oneself and others as intentional agents
 - realize that other agents may have different goals

DEVELOPMENTAL TIMELINE AND ARTIFICIAL MODELS

- the following part will discuss the development of the necessary skills throughout childhood
- computational and robotic models of these skills
- the structure will be oriented to the four prerequisites we just discussed
 - attention detection
 - attention manipulation
 - social coordination
 - intentional understanding

ATTENTION DETECTION

- 0-3 months: children exhibit strong preference for looking at faces
- 6 months: look to the right side of the room, following an adults view
- 9 months: gaze direction can be *accurately* detected
- 12 months: the correct object can be attended to
 - vergence and probably context play a role
 - only in the field of view
- 18 months: objects outside the field of view

MODELS OF ATTENTION DETECTION

- some basic skills have been implemented
 - face detection, eye-contact
- some researchers focus on modeling the **development** of attention detection instead of the skills
 - robots can learn to interpret each others pointing gestures
 - to follow human gaze in order to find objects
- modeling attention detection is the best investigated area so far

ATTENTION MANIPULATION

- distinction between drawing attention to oneself and to others/objects
- child's control during mutual gaze is limited first and extends over time
- 9 months: *imperative* pointing occurs
 - a request for an object using a gesture
 - occurs regardless of someone else's presence in the room

ATTENTION MANIPULATION

(2)

- 12 months: *declarative* pointing
 - emerges shortly before the use of verbal symbols
 - it is clearly used to draw someones attention to something
 - it is still an open debate wether the pointing skill results from imitation or from capabilities in attention following
- 18 months: first predications
 - child points at something which is to be the topic
 - verbal comments to draw attention towards some aspect of it

MODELS OF ATTENTION MANIPULATION

- several issues in the development of attention manipulation haven't been addressed yet
 - How can pointing emerge from grasping behavior?
 - How does declarative pointing appear?
 - By which processes can words replace gestures for drawing attention?
- Existing model: Robovie is able to attract attention by establishing mutual gaze and pointing at an object
 - <http://www.youtube.com/watch?v=n0z0ZK5MEbU>

SOCIAL COORDINATION

- social coordination is a crucial prerequisite for the development of social cognition
- protoconversation
 - six-week old children play games with their caregiver
 - adults treat these early social responses as normal social behavior
- shared **routines**
 - each caregiver develops their own set of conventional games
 - ritualized structures are crucial to learn roles in social exchanges (Kaye, 1982)
 - the routines are the result of a negotiation process

SOCIAL COORDINATION(2)

- Joint activity and imitative games
 - 9 months: complex forms of coordinated activity
 - imitation is crucial for learning social exchanges for turn-taking and role-switching
 - goal sharing starts to emerge (Tomasello 2004)
- coordination of action plans
 - collaborations and organization of role-switches
 - complex imitative games, first verbal exchanges

MODELS OF SOCIAL COORDINATION

- many works have emphasized importance of structured interactions for the development of higher social skills
- yet there are hardly any models of the development of shared interactional routines
 - robots capable of learning turn-taking via evolutionary algorithms (Ikegami and Iizuka, 2003)
- experiments of the emergence of coordinated behavior from tasks that involve cooperation and coordination
 - Quinn (2001, 2003): evolved a team of mobile robots for the ability to move by remaining close to one another
 - <http://www.youtube.com/watch?v=SkvpEfAPXn4&feature=fvw>

MODELS OF IMITATION

- HRI-approaches
 - robots that immediately imitate simple motor skills
 - this imitation however is not investigated in it's role for social coordination
- As in most of the other fields, lots of issues remain
 - What kind of reward structure is needed for interaction to emerge?
 - What dynamics lead to the formation of turns during interaction?
 - How is the structure of new games captured?

INTENTIONAL UNDERSTANDING

- Tomasello argues that intentional understanding emerges at 12 months
- this leads to a qualitative shift on all levels of behavior
- there are different theories concerning the shift
 - nativist: children view autonomous behavior as intentional from birth
 - culturalist: it is totally learned from experience

NECESSARY SKILLS

- the necessary skills are hard to identify and thus a point of heavy debate
- parsing skills
 - find regularities in behavior
 - segment observed behavior to action units
 - find reliable perceptual features of action units
- make inferences and plans about hidden states

IMPORTANT DEVELOPMENTAL STEPS

- 7 months: distinction between animate/inanimate as well as social/physical causality
- 12 months: causal links between gaze direction and intentional action
- 18 months: experimental evidence for intentional understanding multiplies

MODELING INTENTIONAL UNDERSTANDING

- goals and intentions have always been important notions in AI
- behavior parsing remains an open issue
- problem of 'what to imitate'
- this area is hardest to investigate, so far no real progress

INTEGRATION OF THE SKILLS

- What drives the four interrelated skills?
 - Tomasello: two ontogenetic pathways
- as has been shown most of the models focus on single skills
- synergetic dynamics might be missed
- this is why many researchers want to go into the direction of holistic modeling

CONCLUSIONS

- as far as modeling is concerned the task seems to be a fragmented puzzle
 - lots of work on attention detection
 - intentional understanding is most underinvestigated
- one important line of investigation is imitation
- it remains an open question if robots will someday be capable to engage in joint attention with another agent

THANKS FOR YOUR
ATTENTION!

Questions?