Natural Machines

Speaker: Benjamin Hollinger
Chair: Quan Nguyen
Supervisor: Patrick Gebhard

Advances in Human-Computer Interaction
WS 2007/08
Contents

• Introduction
• Classes of social robots
• Methodology
• Realizations
  − Kismet
  − COG
• Conclusions
Introduction

- Growing number of service robots
Introduction

- Growing number of service robots
- Tasks become more and more complex
Introduction

- Growing number of service robots
- Tasks become more and more complex
  => Robots become more and more complex
Introduction

- Growing number of service robots
- Tasks become more and more complex
  => Robots become more and more complex
- How can untrained users interface with these sophisticated devices in an
Introduction

- Growing number of service robots
- Tasks become more and more complex
  => Robots become more and more complex
- How can untrained users interface with these sophisticated devices in an
  - intuitive,
Introduction

• Growing number of service robots
• Tasks become more and more complex
  => Robots become more and more complex
• How can untrained users interface with these sophisticated devices in an
  - intuitive,
  - efficient
Introduction

- Growing number of service robots
- Tasks become more and more complex
  => Robots become more and more complex
- How can untrained users interface with these sophisticated devices in an
  - intuitive,
  - efficient and
  - enjoyable way?
Introduction

• Growing number of service robots
• Tasks become more and more complex
  => Robots become more and more complex
• How can untrained users interface with these sophisticated devices in an
  − intuitive,
  − efficient and
  − enjoyable way?
• Answer: Socially intelligent robots.
Classes of social robots

- Socially evocative
  - Easy to anthropomorphize
  - Evoke feelings in users
Classes of social robots

- **Socially evocative**
  - Easy to anthropomorphize
  - Evoke feelings in users
- **Social interface**
  - Provide a natural interface
  - No deep cognition model required
Classes of social robots

- Socially evocative
  - Easy to anthropomorphize
  - Evoke feelings in users
- Social interface
  - Provide a natural interface
  - No deep cognition model required
- Socially receptive
  - Passive as social actors
  - Can benefit from social interactions
  - Learn by imitation
Classes of social robots (cont.)

- Sociable
  - Pro-active social actors
  - Possess social goals, drives and emotions
  - Usually incorporate a deep model of social cognition.
Methodology - Goals

- Human-oriented perception
  - Detect and interpret human behaviour
  - Recognizing gestures, activity and intent
Methodology - Goals

- Human-oriented perception
  - Detect and interpret human behaviour
  - Recognizing gestures, activity and intent
- Natural human-robot interaction
  - Establish appropriate social expectations
  - Follow social conventions and norms
Methodology - Goals

- **Human-oriented perception**
  - Detect and interpret human behaviour
  - Recognizing gestures, activity and intent

- **Natural human-robot interaction**
  - Establish appropriate social expectations
  - Follow social conventions and norms

- **Readable social cues**
  - Provide feedback of its emotional state
  - Intuitive, transparent interface
Methodology - Goals

- Human-oriented perception
  - Detect and interpret human behaviour
  - Recognizing gestures, activity and intent
- Natural human-robot interaction
  - Establish appropriate social expectations
  - Follow social conventions and norms
- Readable social cues
  - Provide feedback of its emotional state
  - Intuitive, transparent interface
- Real-time performance
  - On par with human performance
  - No “loading times”
Methodology – Primary Designs

- Biologically inspired
  - Designed after biological systems
  - Implements drives, emotions, behaviour and motor systems as found in nature
  - Realizes believable social behaviour from “ground up”
Methodology – Primary Designs

- **Biologically inspired**
  - Designed after biological systems
  - Implements drives, emotions, behaviour and motor systems as found in nature
  - Realizes believable social behaviour from “ground up”

- **Functionally designed**
  - No deeper model for cognition
  - Robot has a function and uses social behaviour mainly as interface
  - “Shallow” cognition
  - Robot is “faking it”
Methodology - Embodiment

- Measured in possible interactions with the world
Methodology - Embodiment

- Measured in possible interactions with the world
- Sets expectations (form suggests function)
Methodology - Embodiment

- Measured in possible interactions with the world
- Sets expectations (form suggests function)
- Different types
Methodology - Embodiment

- Measured in possible interactions with the world
- Sets expectations (form suggests function)
- Different types
  - Anthropomorphic
Methodology - Embodiment

- Measured in possible interactions with the world
- Sets expectations (form suggests function)
- Different types
  - Anthropomorphic
    - Human-like form
Methodology - Embodiment

- Measured in possible interactions with the world
- Sets expectations (form suggests function)
- Different types
  - Anthropomorphic
    - Human-like form
    - Best suited for human-robot social interaction
Methodology - Embodiment

- Measured in possible interactions with the world
- Sets expectations (form suggests function)
- Different types
  - Anthropomorphic
    - Human-like form
    - Best suited for human-robot social interaction
  - Zoomorphic
Methodology - Embodiment

- Measured in possible interactions with the world
- Sets expectations (form suggests function)
- Different types
  - Anthropomorphomorphic
    - Human-like form
    - Best suited for human-robot social interaction
  - Zoomorphomorphic
    - Form imitates animal
Methodology - Embodiment

- Measured in possible interactions with the world
- Sets expectations (form suggests function)
- Different types
  - Anthropomorphic
    - Human-like form
    - Best suited for human-robot social interaction
  - Zoomorphic
    - Form imitates animal
    - Indicates pet-owner relation
Methodology - Embodiment

- Measured in possible interactions with the world
- Sets expectations (form suggests function)
- Different types
  - Anthropomorphic
    - Human-like form
    - Best suited for human-robot social interaction
  - Zoomorphic
    - Form imitates animal
    - Indicates pet-owner relation
    - Makes peer-to-peer interaction difficult
Methodology - Embodiment

- Caricatured
  - Exaggerates one or more physical feature
  - Good for implying a certain function or ability
  - Makes it easy to interact with in a playful manner
  - No danger of falling into the “Uncanny Valley” [1]

Methodology - Embodiment

- Caricatured
  - Exaggerates one or more physical feature
  - Good for implying a certain function or ability
  - Makes it easy to interact with in a playful manner
  - No danger of falling into the “Uncanny Valley” [1]

Methodology - Embodiment

- **Caricatured**
  - Exaggerates one or more physical feature
  - Good for implying a certain function or ability
  - Makes it easy to interact with in a playful manner
  - No danger of falling into the “Uncanny Valley” [1]

Methodology - Embodiment

- Caricatured
  - Exaggerates one or more physical feature
  - Good for implying a certain function or ability
  - Makes it easy to interact with in a playful manner
  - No danger of falling into the “Uncanny Valley” [1]

Methodology - Embodiment

- Caricatured
  - Exaggerates one or more physical feature
  - Good for implying a certain function or ability
  - Makes it easy to interact with in a playful manner
  - No danger of falling into the “Uncanny Valley” [1]

Methodology - Embodiment

- **Caricatured**
  - Exaggerates one or more physical feature
  - Good for implying a certain function or ability
  - Makes it easy to interact with in a playful manner
  - No danger of falling into the “Uncanny Valley” [1]

- **Functional**
  - Function dictates form
  - No realistic form desired

Methodology - Personality

- Expressed via the “Big Five” model
Methodology - Personality

- Expressed via the “Big Five” model
  - Agreeableness
  - Conscientiousness
  - Extroversion
  - Openness
  - Neuroticism
Methodology - Personality

- Expressed via the “Big Five” model
  - Agreeableness
  - Conscientiousness
  - Extroversion
  - Openness
  - Neuroticism

- Five common archetypes:
Methodology - Personality

- Expressed via the “Big Five” model
  - Agreeableness
  - Conscientiousness
  - Extroversion
  - Openness
  - Neuroticism

- Five common archetypes:
  - Tool-like
Methodology - Personality

- Expressed via the “Big Five” model
  - Agreeableness
  - Conscientiousness
  - Extroversion
  - Openness
  - Neuroticism

- Five common archetypes:
  - Tool-like
  - Pet or Creature
Methodology - Personality

- Expressed via the “Big Five” model
  - Agreeableness
  - Conscientiousness
  - Extroversion
  - Openness
  - Neuroticism

- Five common archetypes:
  - Tool-like
  - Pet or Creature
  - Cartoon
Methodology - Personality

- Expressed via the “Big Five” model
  - Agreeableness
  - Conscientiousness
  - Extroversion
  - Openness
  - Neuroticism

- Five common archetypes:
  - Tool-like
  - Pet or Creature
  - Cartoon
  - Artificial Being
Methodology - Personality

- Expressed via the “Big Five” model
  - Agreeableness
  - Conscientiousness
  - Extroversion
  - Openness
  - Neuroticism

- Five common archetypes:
  - Tool-like
  - Pet or Creature
  - Cartoon
  - Artificial Being
  - Human-like
Methodology - Emotions

- Described as
Methodology - Emotions

• Described as
  • Discrete categories (OCC-Model [1], Ekman's Basics[2])
  • A point in a dimensional space (Valence, Arousal)
  • Both

Methodology - Emotions

- Described as
  - Discrete categories (OCC-Model \[1\], Ekman's Basics\[2\])
  - A point in a dimensional space (Valence, Arousal)
  - Both

- Communicated via

\[1\] Ortony, Clore, Collins 1988
\[2\] Ekman 1972
Methodology - Emotions

- Described as
  - Discrete categories (OCC-Model \textsuperscript{[1]}, Ekman's Basics\textsuperscript{[2]})
  - A point in a dimensional space (Valence, Arousal)
  - Both

- Communicated via
  - Speech
  - Facial Expression
  - Body Language

\textsuperscript{[1]} Ortony, Clore, Collins 1988
\textsuperscript{[2]} Ekman 1972
Methodology - Emotions

- Described as
  - Discrete categories (OCC-Model [1], Ekman's Basics[2])
  - A point in a dimensional space (Valence, Arousal)
  - Both

- Communicated via
  - Speech
  - Facial Expression
  - Body Language

- Used as

Methodology - Emotions

- Described as
  - Discrete categories (OCC-Model \[1\], Ekman's Basics\[2\])
  - A point in a dimensional space (Valence, Arousal)
  - Both

- Communicated via
  - Speech
  - Facial Expression
  - Body Language

- Used as
  - Control mechanism for behaviour and feedback
  - Different behaviour in different moods (sad, happy)

\[1\] Ortony, Clore, Collins 1988
\[2\] Ekman 1972
Methodology - Communication

- Low-level
Methodology - Communication

- Low-level
  - Simple / synthetic words or symbols
  - utterances
Methodology - Communication

- Low-level
  - Simple / synthetic words or symbols
  - utterances
- Non-verbal
Methodology - Communication

- Low-level
  - Simple / synthetic words or symbols
  - utterances
- Non-verbal
  - Body language
  - Physical actions to communicate (Pointing,...)
  - Is a useful alternative
Methodology - Communication

- **Low-level**
  - Simple / synthetic words or symbols
  - utterances

- **Non-verbal**
  - Body language
  - Physical actions to communicate (Pointing,...)
  - Is a useful alternative

- **Natural Language**
Methodology - Communication

- **Low-level**
  - Simple / synthetic words or symbols
  - utterances

- **Non-verbal**
  - Body language
  - Physical actions to communicate (Pointing,...)
  - Is a useful alternative

- **Natural Language**
  - High-level speech
  - Clearly the most intuitive
  - Hard to realize
Methodology - Perception

• Conventional goals:
Methodology - Perception

- Conventional goals:
  - Localization
  - Navigation
  - Obstacle avoidance
Methodology - Perception

• Conventional goals:
  • Localization
  • Navigation
  • Obstacle avoidance

• Interaction with humans requires a similar perception of the world around them
Methodology - Perception

- Conventional goals:
  - Localization
  - Navigation
  - Obstacle avoidance

- Interaction with humans requires a similar perception of the world around them

- Human-oriented goals:
Methodology - Perception

- Conventional goals:
  - Localization
  - Navigation
  - Obstacle avoidance
- Interaction with humans requires a similar perception of the world around them
- Human-oriented goals:
  - Feature-tracking (Face, Body, Hands)
Methodology - Perception

- **Conventional goals:**
  - Localization
  - Navigation
  - Obstacle avoidance

- **Interaction with humans requires a similar perception of the world around them**

- **Human-oriented goals:**
  - Feature-tracking (Face, Body, Hands)
  - Feature-interpretation (Gestures, facial expressions,..)
Methodology - Perception

- Conventional goals:
  - Localization
  - Navigation
  - Obstacle avoidance

- Interaction with humans requires a similar perception of the world around them

- Human-oriented goals:
  - Feature-tracking (Face, Body, Hands)
  - Feature-interpretation (Gestures, facial expressions,..)
  - Speech-interpretation (affective, commands,...)
Methodology (cont.)

- User modeling
Methodology (cont.)

- User modeling
  - Enables more sophisticated reaction to user and user behavior
Methodology (cont.)

- User modeling
  - Enables more sophisticated reaction to user and user behavior
  - Several methods of user modeling (quantitative, qualitative, static, dynamic)
Methodology (cont.)

• User modeling
  − Enables more sophisticated reaction to user and user behavior
  − Several methods of user modeling (quantitative, qualitative, static, dynamic)

• Intentionality
Methodology (cont.)

- **User modeling**
  - Enables more sophisticated reaction to user and user behavior
  - Several methods of user modeling (quantitative, qualitative, static, dynamic)
- **Intentionality**
  - Attention (shared attention, object tracking)
Methodology (cont.)

• User modeling
  – Enables more sophisticated reaction to user and user behavior
  – Several methods of user modeling (quantitative, qualitative, static, dynamic)

• Intentionality
  – Attention (shared attention, object tracking)
  – Expression (goal-directed behaviour,...)
Methodology (cont.)

- **User modeling**
  - Enables more sophisticated reaction to user and user behavior
  - Several methods of user modeling (quantitative, qualitative, static, dynamic)

- **Intentionality**
  - Attention (shared attention, object tracking)
  - Expression (goal-directed behaviour,...)

- **Socially situated learning**
Methodology (cont.)

- **User modeling**
  - Enables more sophisticated reaction to user and user behavior
  - Several methods of user modeling (quantitative, qualitative, static, dynamic)

- **Intentionality**
  - Attention (shared attention, object tracking)
  - Expression (goal-directed behaviour,...)

- **Socially situated learning**
  - Most common: Imitation
Methodology (cont.)

- **User modeling**
  - Enables more sophisticated reaction to user and user behavior
  - Several methods of user modeling (quantitative, qualitative, static, dynamic)

- **Intentionality**
  - Attention (shared attention, object tracking)
  - Expression (goal-directed behaviour, ...)

- **Socially situated learning**
  - Most common: Imitation
  - Open questions (When? What? How to integrate? How to evaluate?)
Realizations - Kismet
Kismet - Overview

- Developed at MIT
Kismet - Overview

- Developed at MIT
- Sociable Robot in the role of an infant with
Kismet - Overview

- Developed at MIT
- Sociable Robot in the role of an infant with
  - Models for
    - Drives
    - Emotions
    - Behaviour
Kismet - Overview

- Developed at MIT
- Sociable Robot in the role of an infant with
  - Models for
    - Drives
    - Emotions
    - Behaviour
  - The ability to communicate its current emotional state via facial expressions.
Kismet - Drives

- Kismet is driven by three motivations:
Kismet - Drives

- Kismet is driven by three motivations:
  - Social drive (Interact with people)
Kismet - Drives

• Kismet is driven by three motivations:
  – Social drive (Interact with people)
  – Stimulation drive (Play with toys)
Kismet - Drives

- Kismet is driven by three motivations:
  - Social drive (Interact with people)
  - Stimulation drive (Play with toys)
  - Fatigue drive (Rest)
Kismet - Drives

• Kismet is driven by three motivations:
  − Social drive (Interact with people)
  − Stimulation drive (Play with toys)
  − Fatigue drive (Rest)

• A drive can either be in a
Kismet - Drives

• Kismet is driven by three motivations:
  - Social drive (Interact with people)
  - Stimulation drive (Play with toys)
  - Fatigue drive (Rest)

• A drive can either be in a
  - Overwhelmed state
Kismet - Drives

• Kismet is driven by three motivations:
  − Social drive (Interact with people)
  − Stimulation drive (Play with toys)
  − Fatigue drive (Rest)

• A drive can either be in a
  − Overwhelmed state
  − Homeostatic state
Kismet - Drives

- Kismet is driven by three motivations:
  - Social drive (Interact with people)
  - Stimulation drive (Play with toys)
  - Fatigue drive (Rest)

- A drive can either be in a
  - Overwhelmed state
  - Homeostatic state
  - Under-stimulated state
Kismet - Drives

- Kismet is driven by three motivations:
  - Social drive (Interact with people)
  - Stimulation drive (Play with toys)
  - Fatigue drive (Rest)
- A drive can either be in a
  - Overwhelmed state
  - Homeostatic state
  - Under-stimulated state
- Directly affect emotions and behaviour.
Kismet - Emotions

- Emotional state is represented as a point in an affect space along three dimensions

[1]
Kismet - Emotions

- Emotional state is represented as a point in an affect space along three dimensions.
- Affect space is divided into regions.
Kismet - Emotions

- Emotional state is represented as a point in an affect space \(^{[1]}\) along three dimensions
- Affect space is divided into regions
- Position in affect space directly influenced by drive saturation.

\[^{1}\] Breazeal 1999
Kismet - Behaviour

- Chosen behaviour influenced by
Kismet - Behaviour

- Chosen behaviour influenced by
  - Drive saturation
Kismet - Behaviour

- Chosen behaviour influenced by
  - Drive saturation,
  - Emotional state
Kismet - Behaviour

• Chosen behaviour influenced by
  - Drive saturation,
  - Emotional state,
  - Perceived environment
Kismet - Behaviour

- Chosen behaviour influenced by
  - Drive saturation,
  - Emotional state,
  - Perceived environment
- Behaviours are organized in a hierarchy with three levels
Chosen behaviour influenced by
- Drive saturation,
- Emotional state,
- Perceived environment

Behaviours are organized in a hierarchy with three levels:
- Global task level
Kismet - Behaviour

• Chosen behaviour influenced by
  - Drive saturation,
  - Emotional state,
  - Perceived environment

• Behaviours are organized in a hierarchy with three levels:
  - Global task level
  - Strategy level (prioritized by drive and emotional state)
Kismet - Behaviour

• Chosen behaviour influenced by
  − Drive saturation,
  − Emotional state,
  − Perceived environment

• Behaviours are organized in a hierarchy with three levels:
  − Global task level
  − Strategy level (prioritized by drive and emotional state)
  − Sub-Task level (expressive output)
Kismet – Behaviour (cont.)
Kismet – Behaviour (cont.)
Kismet – Behaviour (cont.)
Kismet – Behaviour (cont.)
Kismet – Facial Expressions

- Kismet has a 15 Degree-of-Freedom face
Kismet – Facial Expressions

- Kismet has a 15 Degree-of-Freedom face
  - Eyelids
  - Eyebrows
  - Ears
  - Lips
Kismet – Facial Expressions

- Kismet has a 15 Degree-of-Freedom face
  - Eyelids
  - Eyebrows
  - Ears
  - Lips
Kismet – Facial Expressions

- Kismet has a 15 Degree-of-Freedom face
  - Eyelids
  - Eyebrows
  - Ears
  - Lips
Kismet – Facial Expressions

- Kismet has a 15 Degree-of-Freedom face
  - Eyelids
  - Eyebrows
  - Ears
  - Lips

- Each region in the affect space has a corresponding facial expression.
Kismet – Emotion-to-face
Realizations - COG
COG - Overview

- Developed at MIT
COG - Overview

- Developed at MIT
- Goal: Human likeness
COG - Overview

- Developed at MIT
- Goal: Human likeness
- Some of COG's abilities:
COG - Overview

- Developed at MIT
- Goal: Human likeness
- Some of COG's abilities:
  - Human-like eye movements
COG - Overview

- Developed at MIT
- Goal: Human likeness
- Some of COG's abilities:
  - Human-like eye movements
    - Saccades (rapid eye movement)
    - Moving target tracking
    - Maintaining eye fixation when head or torso are moving
COG - Overview

- Developed at MIT
- Goal: Human likeness
- Some of COG's abilities:
  - Human-like eye movements
    - Saccades (rapid eye movement)
    - Moving target tracking
    - Maintaining eye fixation when head or torso are moving
  - Face and eye detection
COG - Overview

- Developed at MIT
- Goal: Human likeness
- Some of COG's abilities:
  - Human-like eye movements
    - Saccades (rapid eye movement)
    - Moving target tracking
    - Maintaining eye fixation when head or torso are moving
  - Face and eye detection
    - Imitating someone's nodding or head shaking
    - Gaze tracking
COG - Overview

• Developed at MIT
• Goal: Human likeness
• Some of COG's abilities:
  - Human-like eye movements
    • Saccades (rapid eye movement)
    • Moving target tracking
    • Maintaining eye fixation when head or torso are moving
  - Face and eye detection
    • Imitating someone's nodding or head shaking
    • Gaze tracking
  - Incremental learning:
COG - Overview

• Developed at MIT
• Goal: Human likeness
• Some of COG's abilities:
  - Human-like eye movements
    • Saccades (rapid eye movement)
    • Moving target tracking
    • Maintaining eye fixation when head or torso are moving
  - Face and eye detection
    • Imitating someone's nodding or head shaking
    • Gaze tracking
  - Incremental learning:
    • Break tasks down to atomic actions
    • Combine task to more complex tasks
COG - Learning

- How to grasp an object (~ 2000 tries):
• Learning similar in social contexts
• Learning similar in social contexts
• Basic social interaction: Shared attention
COG – Learning (cont.)

- Learning similar in social contexts
- Basic social interaction: Shared attention
  - Breaks down to
Learning similar in social contexts

Basic social interaction: Shared attention
  - Breaks down to
    • Hold gaze
      - Detect face
      - Detect eyes
      - Move head to meet gaze
• Learning similar in social contexts
• Basic social interaction: Shared attention
  - Breaks down to
    • Hold gaze
      - Detect face
      - Detect eyes
      - Move head to meet gaze
    • Follow gaze to object of interest
      - Find angle of gaze
      - Extrapolate object from angle of gaze
      - Move head to object of interest
• Learning similar in social contexts
• Basic social interaction: Shared attention
  - Breaks down to
    • Hold gaze
      - Detect face
      - Detect eyes
      - Move head to meet gaze
    • Follow gaze to object of interest
      - Find angle of gaze
      - Extrapolate object from angle of gaze
      - Move head to object of interest
  - Can be learned incrementally
Conclusions

• The most important design issue is to convey intentionality
  - Peer-to-peer social interaction
  - Mutually shared control
• Kismet successfully implements a model for drives and emotions
• COG's incremental learning ability is a basis for a human-like perception of the world
• Research has still a long way to go