Script generated by TTT

Title: groh: profile1 (16.06.2015)

Date: Tue Jun 16 14:59:41 CEST 2015

Duration: 90:59 min

Pages: 46

Social Gaming / Social Computing SS 2015

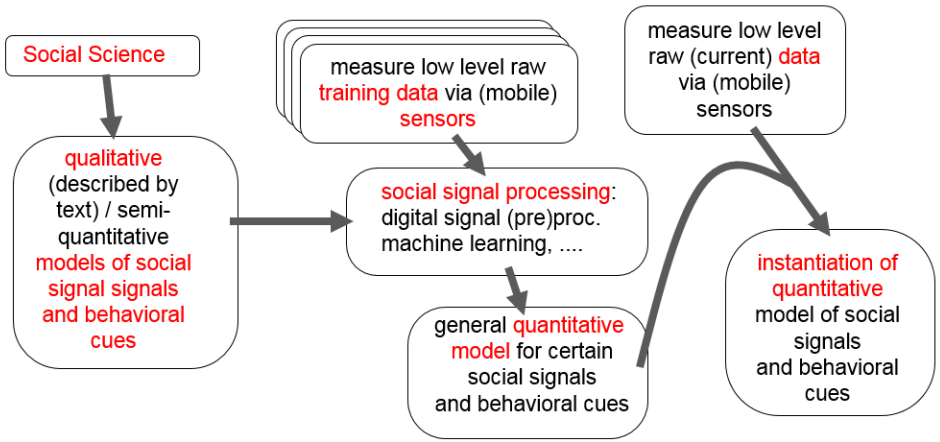
PD Dr. Georg Groh

Applied Informatics / Cooperative Systems AICOS

Technische Universität München

Social Signal Processing

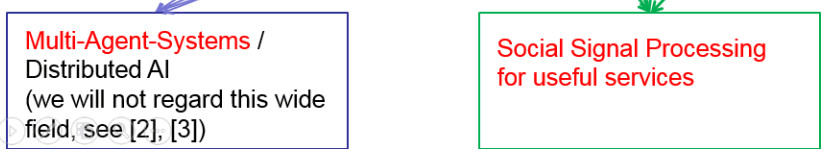
- Up to now: most emphasis on long term social context (e.g. networks of friendships)
- Now: short term social context: social behavior on small spatial and temporal scales: detection + applications



Social Signal Processing

- **Social intelligence** (humans): Ability to express and recognize social signals / social behaviors from other human individuals
  in order to „function" in a society with other human individuals
  in view of (pareto-)optimizing own and other human's
  utility function (survival, reproduction) via cooperation (modified from [1])

- **Social intelligence** (IT systems): Ability to express and recognize social signals / social behaviors from other human and IT-agent individuals
  in order to „function" in a society with other human and IT-agent individuals
  in view of (pareto-)optimizing own and other IT agent's and fellow human's
  utility function (survival, reproduction, ..) via cooperation



Where and how to use SSP techniques for improving services?

- Socially smart services: examples:
  - -- recommender systems:
    use short term *individual* context detected via SSP:
    ++ emotional state → prefer to recommend funny movies when depressed
    (caveat: not social!!)
  - use short term *social* context detected via SSP:
    ++ social situation → use group recommender:
    recommend pareto optimal choices
- etc.

Common name for paradigm: Socially Aware Computing

distinguished from / subclass of: Context Aware Computing
distinguished from: Personalization

Where and how to use SSP techniques for improving services?

- Socially smart services: examples:
  - -- recommender systems:
    use short term *individual* context detected via SSP:
    ++ emotional state → prefer to recommend funny movies when depressed
    (caveat: not social!!)
  - use short term *social* context detected via SSP:
    ++ social situation → use group recommender:
    recommend pareto optimal choices
- etc.

Common name for paradigm: Socially Aware Computing

distinguished from / subclass of: Context Aware Computing
distinguished from: Personalization

Where and how to use SSP techniques for improving services?

- Socially smart services: examples:
  - -- recommender systems:
    use short term *individual* context detected via SSP:
    ++ emotional state → prefer to recommend funny movies when depressed
    (caveat: not social!!)
  - use short term *social* context detected via SSP:
    ++ social situation → use group recommender:
    recommend pareto optimal choices
- etc.

Common name for paradigm: Socially Aware Computing

distinguished from / subclass of: Context Aware Computing
distinguished from: Personalization

SSP and Reality Mining

- Reality Mining: Field of study / application defined by Alex Pentland (MIT)
  - Analyzing all available traces of human behavior (social and also non-social)
  - → derive models for this behavior → scientific knowledge and applications (e.g. prediction)
  - special focus on „non-virtual“, „direct“, „physical“ behaviors recordable via mobile sensors (calling patterns, location, acceleration patterns etc.)
  - not excluded but not in main focus: large networks arising from long term social relations, analysis of communication content etc.
- Reality Mining may use SSP techniques

Social Signals and Behavioral Cues

- Examples of Social Behavior: Expressing attitude towards elements of a social setting:

  -- Mirroring (if mutual attraction)
  -- aggressive turn taking behavior
  -- expression diapproval of sth. (e.g. via disapproving looks)
  -- expression of sympathy / empathy

- Examples for Behavioral Cues:

  -- facial expressions
  -- body posture / interaction geometry
  -- gestures
  -- expressives (laughter etc.)
  -- emotions reflected in speech prosody (rhythm, intonation, stress)

Social Signals and Behavioral Cues

- Examples of Social Behavior: Expressing attitude towards elements of a social setting:

  -- Mirroring (if mutual attraction)
  -- aggressive turn taking behavior
  -- expression diapproval of sth. (e.g. via disapproving looks)
  -- expression of sympathy / empathy

- Examples for Behavioral Cues:

  -- facial expressions
  -- body posture / interaction geometry
  -- gestures
  -- expressives (laughter etc.)
  -- emotions reflected in speech prosody (rhythm, intonation, stress)

Social Signals and Behavioral Cues

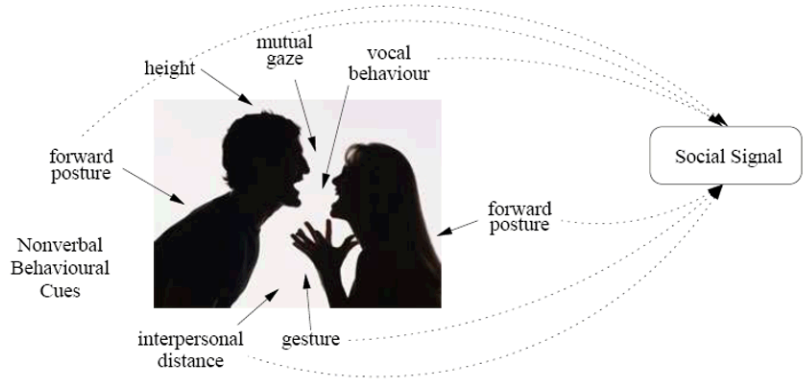


Fig. 1. Behavioural cues and social signals. Multiple behavioural cues (vocal behaviour, posture, mutual gaze, interpersonal distance, etc.) combine to produce a social signal (in this case aggressivity or disagreement) that is evident even if the picture shows only the silhouettes of the individuals involved in the interaction.

[1]

Social Signals and Behavioral Cues

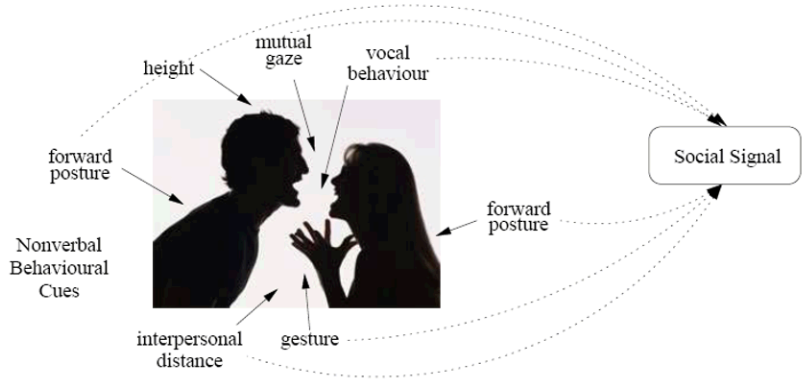


Fig. 1. Behavioural cues and social signals. Multiple behavioural cues (vocal behaviour, posture, mutual gaze, interpersonal distance, etc.) combine to produce a social signal (in this case aggressivity or disagreement) that is evident even if the picture shows only the silhouettes of the individuals involved in the interaction.

[1]

Social Signals and Behavioral Cues

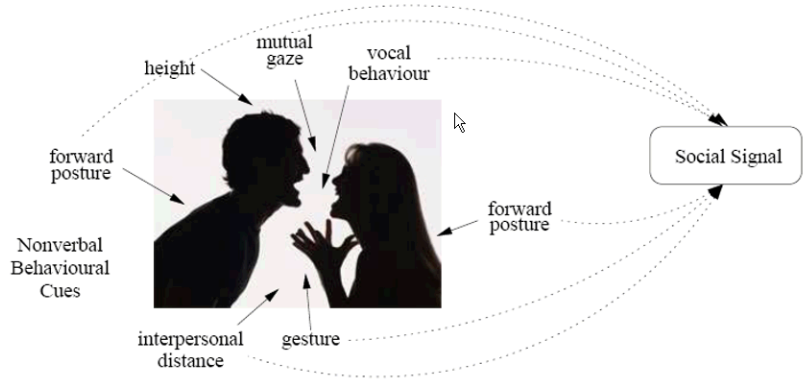


Fig. 1. Behavioural cues and social signals. Multiple behavioural cues (vocal behaviour, posture, mutual gaze, interpersonal distance, etc.) combine to produce a social signal (in this case aggressivity or disagreement) that is evident even if the picture shows only the silhouettes of the individuals involved in the interaction.

[1]

Social Signals and Behavioral Cues

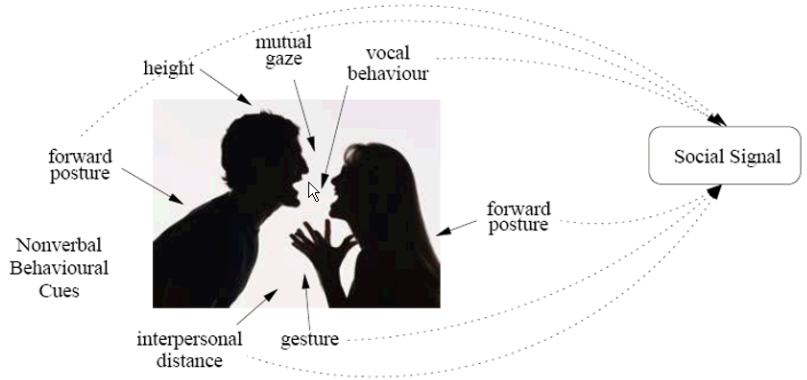


Fig. 1. Behavioural cues and social signals. Multiple behavioural cues (vocal behaviour, posture, mutual gaze, interpersonal distance, etc.) combine to produce a social signal (in this case aggressivity or disagreement) that is evident even if the picture shows only the silhouettes of the individuals involved in the interaction.

[1]

Social Signals, Social Behavior, Behavioral Cues

- (Non-verbal) (Social) Behavioral Cues: „composed of“ / manifested via (series of / parallel / overlapping / single ...) time-series of perceivable or measurable, non-verbal physiological activity.
  (neglecting content of communication)
- (Non-verbal) Social Signals (conscious or unconscious): „composed of“ / manifested via (series of / parallel / overlapping / single ...) Behavioral Cues
- (Non-verbal) Social Behavior: „composed of“/ manifested via (series of / parallel / overlapping / single ...) Social Signals.

- Humans: perceived Social Behavior of other humans
  → Social Awareness
- Non-verbal social signalling + behavior: conveys / determines most of social perception of others (compared to e.g. verbal *content* *(only 7 %)*)

Social Signals, Social Behavior, Behavioral Cues

- (Non-verbal) (Social) Behavioral Cues: „composed of“ / manifested via (series of / parallel / overlapping / single ...) time-series of perceivable or measurable, non-verbal physiological activity.
  (neglecting content of communication)
- (Non-verbal) Social Signals (conscious or unconscious): „composed of“ / manifested via (series of / parallel / overlapping / single ...) Behavioral Cues
- (Non-verbal) Social Behavior: „composed of“/ manifested via (series of / parallel / overlapping / single ...) Social Signals.

- Humans: perceived Social Behavior of other humans
  → Social Awareness
- Non-verbal social signalling + behavior: conveys / determines most of social perception of others (compared to e.g. verbal *content* *(only 7 %)*)

Social Signals, Social Behavior, Behavioral Cues

- Temporal domain:
  - Social Signals + Behavioral Cues :
    - typically: microseconds to seconds
  - Social Behavior:
    - minutes to hours or longer

- Types of „messages“ conveyed by Behavioral Cues:

have direct verbal translation, people are aware of

adaptators : involuntary, habitual, honest (e.g. folding arms, crossing away legs)

regulate turn taking and other conversational aspects

- *affective/attitudinal/cognitive states* (e.g. fear, joy, stress, disagreement, ambivalence and inattention),
- *emblems* (culture-specific interactive signals like wink or thumbs up),
- (*manipulators*)(actions used to act on objects in the environment or self-manipulative actions such as lip biting and scratching),
- *illustrators* (actions accompanying speech such as finger pointing and raised eyebrows), and
- *regulators* (conversational mediators such as the exchange of a look, palm pointing, head nods and smiles).

[1]

Social Signals, Social Behavior, Behavioral Cues

- Temporal domain:
  - Social Signals + Behavioral Cues :
    - typically: microseconds to seconds
  - Social Behavior:
    - minutes to hours or longer

- Types of „messages“ conveyed by Behavioral Cues:

have direct verbal translation, people are aware of

adaptators : involuntary, habitual, honest (e.g. folding arms, crossing away legs)

regulate turn taking and other conversational aspects

- *affective/attitudinal/cognitive states* (e.g. fear, joy, stress, disagreement, ambivalence and inattention),
- *emblems* (culture-specific interactive signals like wink or thumbs up),
- (*manipulators*)(actions used to act on objects in the environment or self-manipulative actions such as lip biting and scratching),
- *illustrators* (actions accompanying speech such as finger pointing and raised eyebrows), and
- *regulators* (conversational mediators such as the exchange of a look, palm pointing, head nods and smiles).

[1]

Social Signals, Social Behavior, Behavioral Cues

- Temporal domain:
  - Social Signals + Behavioral Cues :
    - typically: microseconds to seconds
  - Social Behavior:
    - minutes to hours or longer

- Types of „messages“ conveyed by Behavioral Cues:

have direct verbal translation, people are aware of

adaptators : involuntary, habitual, honest (e.g. folding arms, crossing away legs)

regulate turn taking and other conversational aspects

- *affective/attitudinal/cognitive states* (e.g. fear, joy, stress, disagreement, ambivalence and inattention),
- *emblems* (culture-specific interactive signals like wink or thumbs up),
- (*manipulators*)(actions used to act on objects in the environment or self-manipulative actions such as lip biting and scratching),
- *illustrators* (actions accompanying speech such as finger pointing and raised eyebrows), and
- *regulators* (conversational mediators such as the exchange of a look, palm pointing, head nods and smiles).

[1]

Social Signals, Social Behavior, Behavioral Cues

- Table of Behavioral cues:

Behavioral Cues:	Example Social Behaviours (aspects of) emotion	personality	status	dominance	persuasion	regulation	rapport (*)	Tech. speech anlysis	computer vision	biometry
Physical appearance										
height			√	√					√	√
attractiveness		√	√	√	√		√		√	√
body shape		√		√					√	√
Gesture and posture										
hand gestures	√	√			√	√	√		√	√
posture	√	√	√	√	√	√	√		√	√
walking		√	√	√					√	√
Face and eyes behaviour										
facial expressions	√	√	√	√	√	√	√		√	√
gaze behaviour	√	√	√	√	√	√	√		√	
focus of attention	√	√	√	√	√	√	√		√	
Vocal behaviour										
prosody	√	√		√	√		√	√		
turn taking	√	√	√	√		√	√	√		
vocal outbursts	√	√		√	√	√	√	√		
silence	√		√				√	√		
Space and Environment										
distance	√	√	√		√		√		√	
seating arrangement				√	√		√		√	

(*) mutual trust vs. distrust

[1]

Social Signals, Social Behavior, Behavioral Cues

- Table of Behavioral cues:

Behavioral Cues:	Example Social Behaviours (aspects of): emotion	personality	status	dominance	persuasion	regulation	rapport (*)	Tech.: speech anlysis	computer vision	biometry
Physical appearance										
height			√	√					√	√
attractiveness		√	√	√	√		√		√	√
body shape		√		√					√	√
Gesture and posture										
hand gestures	√	√			√	√	√		√	√
posture	√	√	√	√	√	√	√		√	√
walking		√	√	√					√	√
Face and eyes behaviour										
facial expressions	√	√	√	√	√	√	√		√	√
gaze behaviour	√	√	√	√	√	√	√		√	
focus of attention	√	√	√	√	√	√	√		√	
Vocal behaviour										
prosody	√	√		√	√		√	√		
turn taking	√	√	√	√		√	√	√		
vocal outbursts	√	√		√	√	√	√	√		
silence	√		√				√	√		
Space and Environment										
distance	√	√	√		√		√		√	
seating arrangement				√	√		√		√	

(*) mutual trust vs. distrust

[1]

Social Signals, Social Behavior, Behavioral Cues

- Table of Behavioral cues:

(*) mutual trust vs. distrust

[1]

Physical Appearance

- „Implicit“ „Behavioral“ Cues / Social Signals
  - natural characteristics (e.g. body shape)
  - artifical characteristics (e.g. make up)
- Especially important for attractiveness → halo effect (“what is beautiful is good”);
- other example of effect mechanism: height, sematotype → power, influence, strength

Fig. 2. Somatotypes. The figure shows the three body shapes that tend to elicit the perception of specific personality traits. [1]

Gestures and Posture

- Gestures and posture are related to e.g. expressed emotions (examples: inclination of head, touching own head (various), posture shifts →shame, embarrassment)
  - 90% of gestures : associated with speech (social signals / social signal aspects: illustrators, emblems, regulators etc.)
- unconscious gestures / posture : „honest“ signals → allow to deduce „actual“ / „true“ state / social attitude / etc.
  - Example: adaptators: unconscious expressions of internal states or social attitudes, e.g. manipulation of own body parts (e.g. hair twisting), or objects (e.g. playing with pens); protective gestures (folding arms, rythmically moving legs etc.)

Gestures and Posture

- Gestures and posture are related to e.g. expressed emotions (examples: inclination of head, touching own head (various), posture shifts →shame, embarrassment)
  - 90% of gestures : associated with speech (social signals / social signal aspects: illustrators, emblems, regulators etc.)
- unconscious gestures / posture : „honest“ signals → allow to deduce „actual“ / „true“ state / social attitude / etc.
  - Example: adaptators: unconscious expressions of internal states or social attitudes, e.g. manipulation of own body parts (e.g. hair twisting), or objects (e.g. playing with pens); protective gestures (folding arms, rythmically moving legs etc.)

Gestures and Posture

- Posture: very important Behavioral cues:
  considered as among the most reliable cues
  about actual social attitude towards (elements of) a social setting
- (partly overlapping) Classification axes for postures in interactive social settings:
  - Inclusive vs exclusive postures
  - relative body orientation
  - congruent vs incongruent (→ mirroring)

Congruent postures

Non−congruent postures [1]

- Postures in „general“ social environments:
  may reveal sth. about individual state or general social attitude:
  example: walking style (→ status, dominance)

Gestures and Posture

- Posture: very important Behavioral cues:
  considered as among the most reliable cues
  about actual social attitude towards (elements of) a social setting
- (partly overlapping) Classification axes for postures in interactive social settings:
  - Inclusive vs exclusive postures
  - relative body orientation
  - congruent vs incongruent (→ mirroring)

Congruent postures

Non−congruent postures [1]

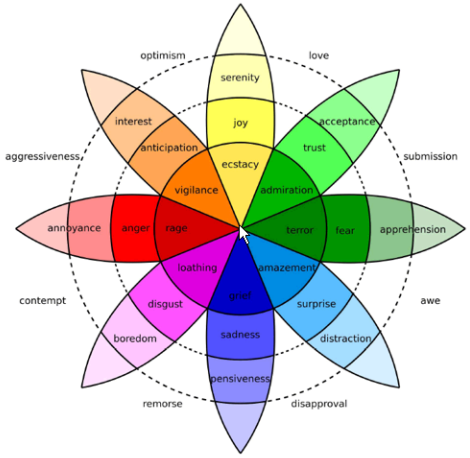
- Postures in „general“ social environments:
  may reveal sth. about individual state or general social attitude:
  example: walking style (→ status, dominance)

Interaction Geometry

- Interaction geometry: special aspects of posture and space and environment:
  relative geometry of interacting persons:
  -- relative body angle
  -- interpersonal distances
  -- relative body positions
- → we will take a more extensive look at a specific example later

Emotion Enconding Systems

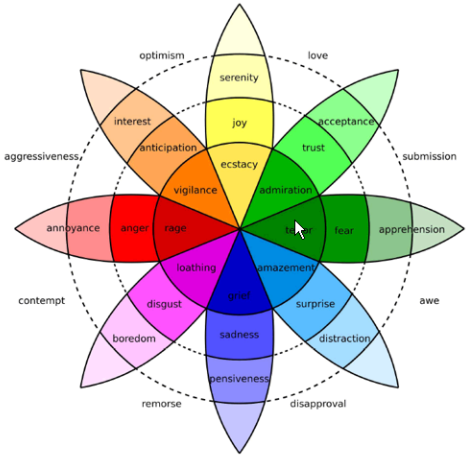
- System to describe (basic) message:
  - 6 basis emotions (Ekman):
    fear, sadness, happiness,
    anger, disgust, surprise
- Plutchik's Wheel of Emotion:



Plutchik's Wheel of Emotion. Source: [Plutchik, 2012; in (1)]

Emotion Enconding Systems

- System to describe (basic) message:
  - 6 basis emotions (Ekman):
    fear, sadness, happiness,
    anger, disgust, surprise
- Plutchik's Wheel of Emotion:



Plutchik's Wheel of Emotion. Source: [Plutchik, 2012; in (1)]

Face and Eye Behavior

- System to describe facial expressions and facial actions:
  Facial Action Coding System (FACS):
  Action Units (AUs):

  Smallest discernable movements of a distinct muscle / muscle group in a face which may take part in facial expressions and facial actions and that may be algorithmically detected

  -- 9 AUs in upper face
  -- 18 AUs in lower face
  -- 11 AUs for head position
  -- 14 additional descriptors

Face and Eye Behavior

- System to describe facial expressions and facial actions:
  Facial Action Coding System (FACS):
  Action Units (AUs):

  Smallest discernable movements of a distinct muscle / muscle group in a face which may take part in facial expressions and facial actions and that may be algorithmically detected

  -- 9 AUs in upper face
  -- 18 AUs in lower face
  -- 11 AUs for head position
  -- 14 additional descriptors

Face and Eye Behavior

- Example studies: FACS patterns → social signals / individual states: recognition of
  - basic emotions
  - cognitive states like interest and puzzlement
  - psychological states like suicidal depression or pain
  - social behavior aspects like rapport (mutual trust vs. distrust),
  - personality traits like extraversion and temperament
  - social signals : e.g. status, trustworthiness,
    emblems (wink, thumbs up, etc.),
    regulators (conversational mediators like nod and gaze exchange)
    illustrators (accompanying speech (e.g. raised eyebrows))

Face and Eye Behavior

- Example studies: FACS patterns → social signals / individual states: recognition of
  - basic emotions
  - cognitive states like interest and puzzlement
  - psychological states like suicidal depression or pain
  - social behavior aspects like rapport (mutual trust vs. distrust),
  - personality traits like extraversion and temperament
  - social signals : e.g. status, trustworthiness,
    emblems (wink, thumbs up, etc.),
    regulators (conversational mediators like nod and gaze exchange)
    illustrators (accompanying speech (e.g. raised eyebrows))

Face and Eye Behavior

- Interesting studies: FACS + classifiers can be better at distinguishing between deception and intoxication (and also between lying and telling the truth) than humans

Vocal Behavior

- Components:
  - voice quality (Prosody)
  - linguistic vocalizations (Segregates)
  - non-linguistic vocalizations
  - silence patterns
  - turn taking patterns

Vocal Behavior

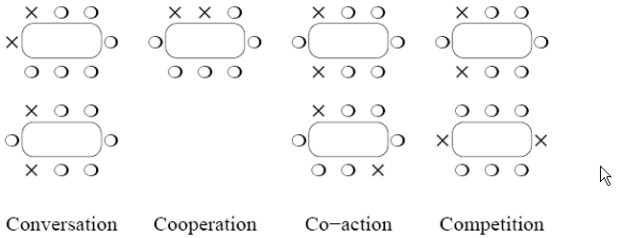
- Silence patterns:
  - -- hesitiation silence: (e.g. ← explaining difficult concepts)
  - -- psycholinguistic silence: ← (language) en-/de-coding difficulties
  - -- interactive silence: expressing respect, doubt, ignoring persons, attract attention to other forms of communication (e.g. gazes)
- Turn-Taking:
  - --regulation of conversations:
    maintaining, yielding, denying or requesting the turn
    ← gaze, voice quality changes, linguistic vocalizations (as backchannelling) at transition relevant points
  - -- coordination of speaker transitions:
    aspects: akward silence, overlapping speech (normal ~ 5-10%

Vocal Behavior

- Silence patterns:
  - -- hesitiation silence: (e.g. ← explaining difficult concepts)
  - -- psycholinguistic silence: ← (language) en-/de-coding difficulties
  - -- interactive silence: expressing respect, doubt, ignoring persons, attract attention to other forms of communication (e.g. gazes)
- Turn-Taking:
  - --regulation of conversations:
    maintaining, yielding, denying or requesting the turn
    ← gaze, voice quality changes, linguistic vocalizations (as backchannelling) at transition relevant points
  - -- coordination of speaker transitions:
    aspects: akward silence, overlapping speech (normal ~ 5-10%

Space and Environment

- Seating arrangements ←→ personality (dominance etc.), social status



Social Signal Processing Chain

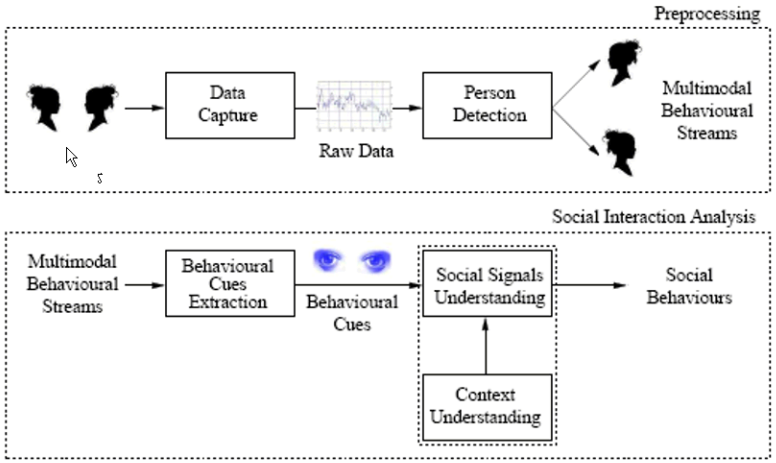


Fig. 6. Machine analysis of social signals and behaviours: a general scheme. The process includes two main stages: The *preprocessing*, takes as input the recordings of social interaction and gives as output the multimodal behavioural streams associated with each person. The *social interaction analysis* maps the multimodal behavioural streams into social signals and social behaviours. [1]

Social Signal Processing Chain

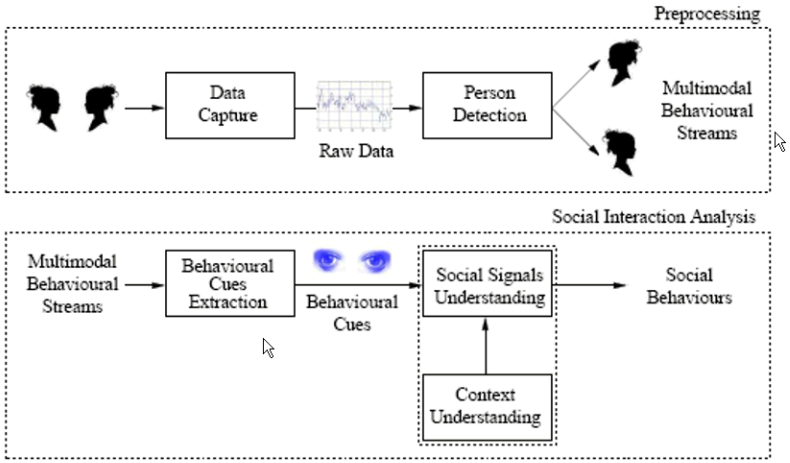


Fig. 6. Machine analysis of social signals and behaviours: a general scheme. The process includes two main stages: The *preprocessing*, takes as input the recordings of social interaction and gives as output the multimodal behavioural streams associated with each person. The *social interaction analysis* maps the multimodal behavioural streams into social signals and social behaviours. [1]

Social Signal Processing Chain

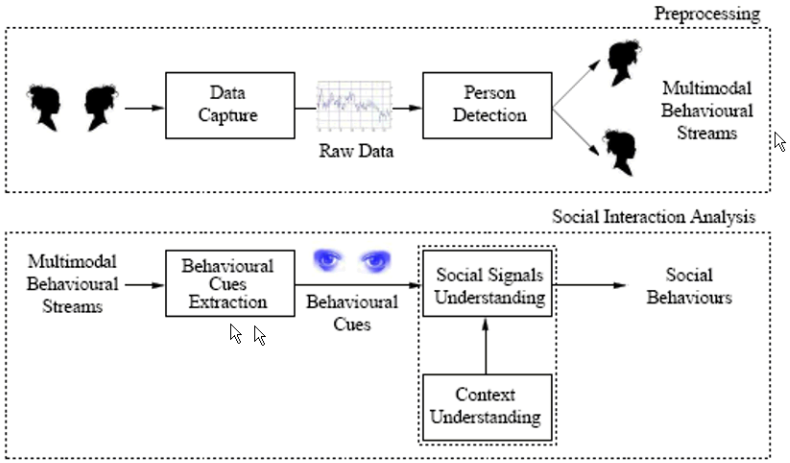


Fig. 6. Machine analysis of social signals and behaviours: a general scheme. The process includes two main stages: The *preprocessing*, takes as input the recordings of social interaction and gives as output the multimodal behavioural streams associated with each person. The *social interaction analysis* maps the multimodal behavioural streams into social signals and social behaviours. [1]

Social Signal Processing Chain

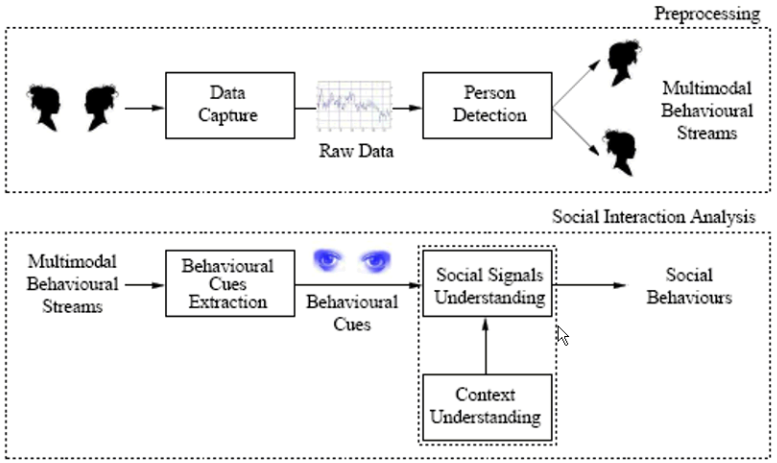


Fig. 6. Machine analysis of social signals and behaviours: a general scheme. The process includes two main stages: The *preprocessing*, takes as input the recordings of social interaction and gives as output the multimodal behavioural streams associated with each person. The *social interaction analysis* maps the multimodal behavioural streams into social signals and social behaviours. [1]

Social Signal Processing Chain

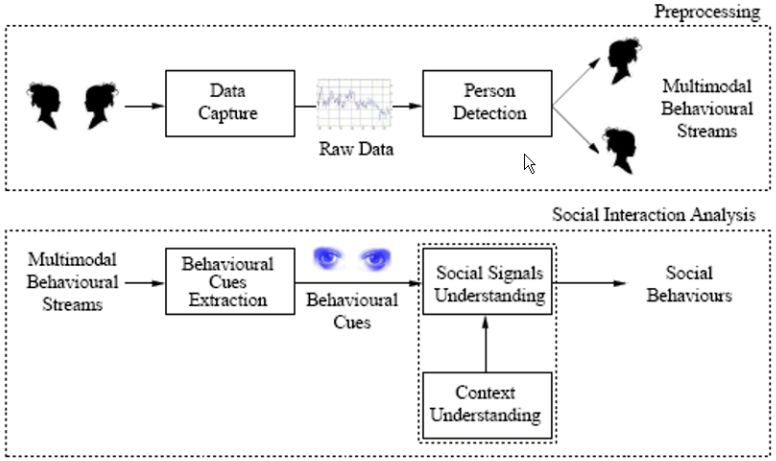


Fig. 6. Machine analysis of social signals and behaviours: a general scheme. The process includes two main stages: The *preprocessing*, takes as input the recordings of social interaction and gives as output the multimodal behavioural streams associated with each person. The *social interaction analysis* maps the multimodal behavioural streams into social signals and social behaviours. [1]

Data Capture

- Infrastructure sensors: (single or multiple) fixed cameras and / or microphones, Kinect devices,
- Mobile sensors: e.g. in mobile phones: GPS, accelerometer, gyroscopes, eye tracker glasses, pulse-meters, EEG devices, etc.
- Issues:
  - Privacy + scientific ethics („principle of informed consent“)
  - passiveness / unobtrousiveness of sensors

Data Capture

- Infrastructure sensors: (single or multiple) fixed cameras and / or microphones, Kinect devices,
- Mobile sensors: e.g. in mobile phones: GPS, accelerometer, gyroscopes, eye tracker glasses, pulse-meters, EEG devices, etc.
- Issues:
  - Privacy + scientific ethics („principle of informed consent“)
  - passiveness / unobtrousiveness of sensors

Person Detection from Audio

- Segmentation into speech / non-speech: (continued)
  -- Use (several) trained binary classifier(s)
  to distinguish between speech and non speech
  on the computed features
- Detection of speaker transitions:
  --split the speech parts into segments (length: e.g. 2-3 seconds)
  --decide with statistical methods, whether
  two segments belong to the same speaker
  (e.g. model each segment with Gaussian,
  use symm. KL-divergence to compute “difference”
  between Gaussians) or
  whether one interval contains one or two speakers
  (e.g. is better modeled by one or two Gaussians, decided
  using likelihood ratio-tests)

Person Detection from Audio

“Speaker Diarization / Segmentation“: given multi-party audio data (possibly with background noise):
→ who talks when?

- Typically 3 steps:
  --segmentation into speech / non-speech
  --detection of speaker transitions
  --clustering of speaker segments (+ classification of speaker)
- Segmentation into speech / non-speech:
  -- Generate features:
  ++ digital signal (pre-) processing
  (involving e.g. sub-division signal into overlapping samples of
  typically several ms, Fourier-transform etc.)
  ++ MEL filters → MEL cepstrum coefficients
  ++ Further Fourier- and other transformations
  ++ additional features: zero-crossing rates, energy statistics etc.