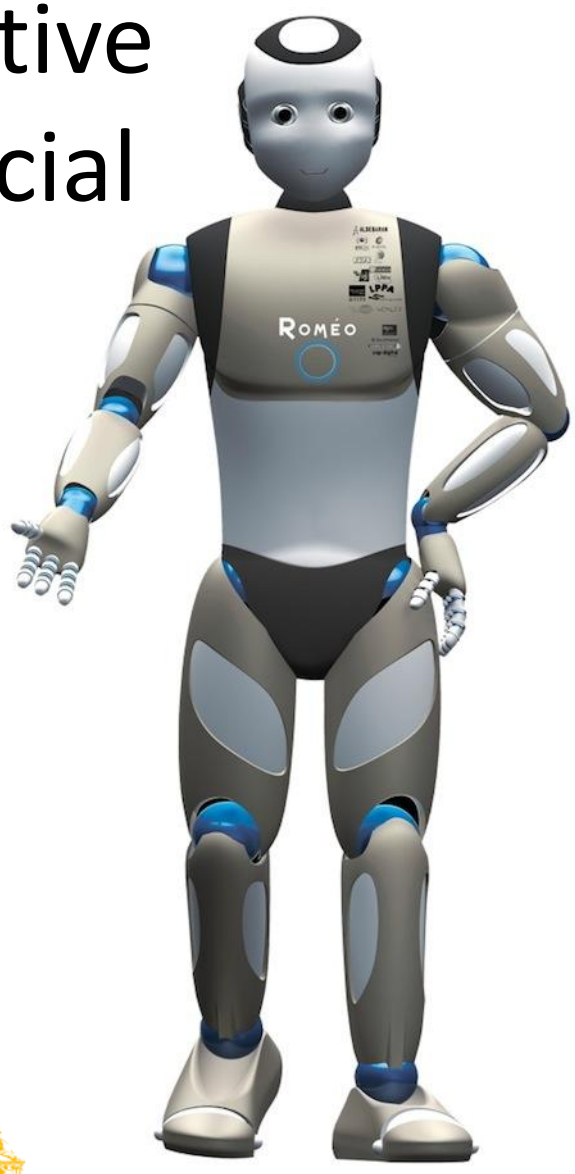
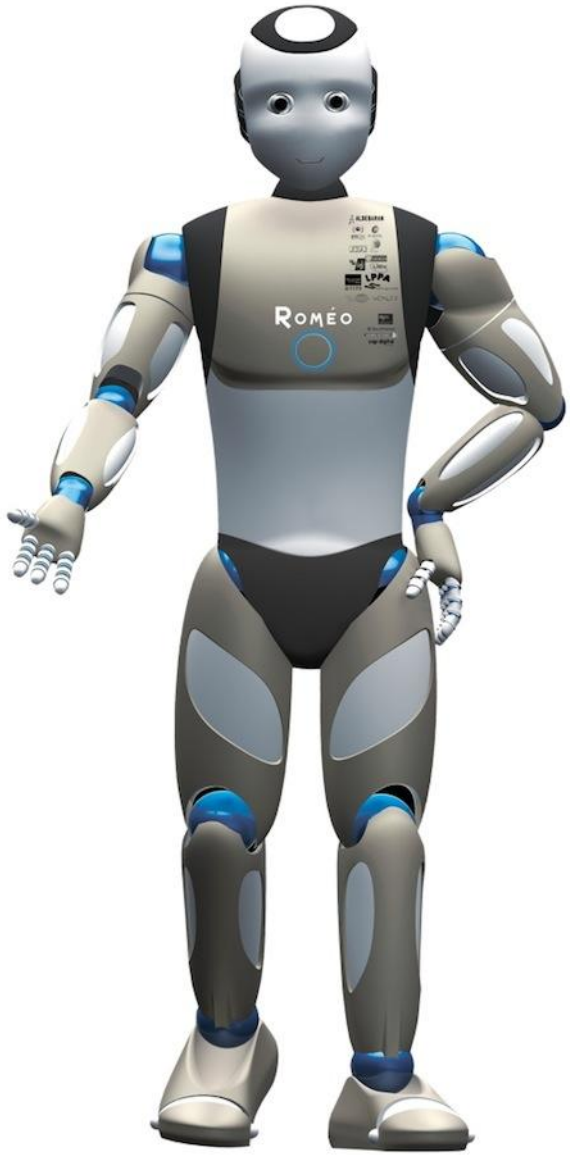
From informative cooperative dialogues to long-term social relation with a robot



Axel Buendia[1] and Laurence Devillers[2]
[1]SPIROPS / CEDRIC CNAM, France
[2]LIMSI-CNRS, University Paris-Sorbonne 4, France

axel.buendia@spirops.com, devil@limsi.fr

The ROMEO project



GOAL

Create a robot to assist elderly people

NEEDS

Socialize

Communicate

Entertain

ALDEBARAN Robotics

SPIROPS
ARTIFICIAL INTELLIGENCE

provides:

→ Tools to edit decisions

→ Tools to debug decisions

→ Emotion models (dynamic & static)

→ Memories management

→ Learning abilities

to

ISIR
INSTITUT DES SYSTÈMES INTELLIGENTS ET DE ROBOTIQUE

CRIIF
www.criif.fr
Centre de Robotique Intégrée d'Ile de France

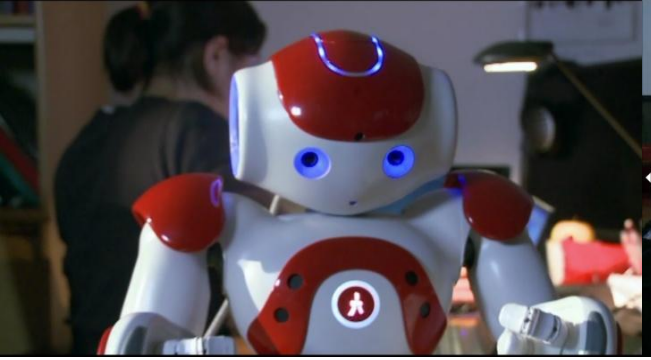
CNRS LIMSI

ALDEBARAN Robotics

Research labs

Robotics companies

to bring life to new robots!



Research at LIMSI-CNRS – Spoken Language G.

Team: « Affective and social dimensions of spoken interactions »

L. Devillers, *C. Chastagnol, A. Delaborde, M. Soury, M. Tahon, C. Vaudable*

http://perso.limsi.fr/Individu/devil/

5 PHD students - Since 2001 team 5-10 researchers

- **“Real-life" emotion perception and detection**: from emotional expression in the voice to multimodal expression up to emotional and mental states in interaction situation
- Applications and projects: robotics, call centers, serious game E-therapy, data-mining

How to realize credible social exchanges and favor the emergence of a social link

- Social links emerge when sharing experiences (social interactions) and memories
- Social interactions require certain abilities such as
  - Social understanding: Planning ahead and dealing with new circumstances.
  - Mind theory: Anticipating the mental state of another person.
- The capacity for deception is necessary for a theory of mind
- Human beings behaviors such as **lies, compassion and jokes** imply that the robot has the ability to represent and understand some complex human beings behaviors.

What about lies, compassion, jokes

- **Lies are used to hide information or to achieve our goals.**
- Deception is a major relational transgression (violates relational rules) that often leads to distrust between relational partners. But lying appears as a normal component of human social interaction
- **Compassion is also of a great importance for social interactions and relationship.**
- Empathy is the capacity to recognize emotions that are being experienced by another. Compassion is useful to read emotions properly and to mirror them.
- **A robot that makes ”jokes” is a matter of context and of correlation with the current subject**.
- It depends also on the type of emotions in the dialogue.

Challenges to the Perception, Dialogue modeling and Artificial Intelligence

- Artificial creature should have the cognitive abilities, the sensing components and the dialoging capabilities to enable it to develop a social behavior
- **Abilities such as anticipation, expectation, memorization and continuous training**
- Such cognitive abilities imply the development of new representations and new AI architecture.

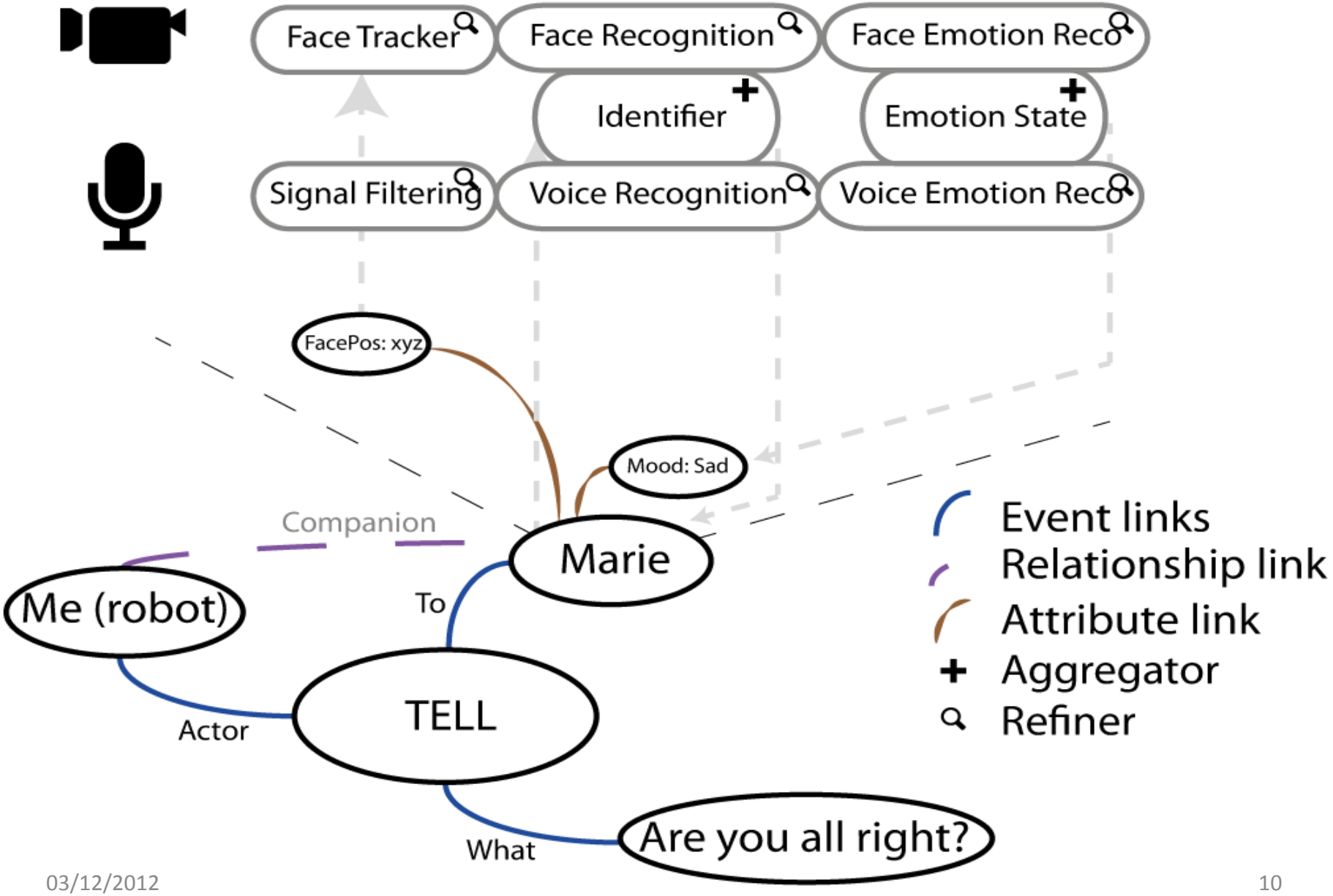
How theories and models in sociology, in psychology, and neurosciences can help

- To develop a long-term human-robot interaction theory is essential for a robotic social companion
- It should be inspired by empirical groundings but also theories and models in :
  - **Sociology (Erving Goffman)** : concepts such as the face, the frames, the social roles, the goals and the working acceptance seem really important for a better model of the social interactions
  - **Psychology (Klaus Scherer)** : component process model of affective states for the appraisal of emotional events
  - **Neurosciences (Vernon Mountcastle)** : how the neo cortical brain is architectured: a bottom-up to extract information from signals and a top-down to create anticipation and expectation
- The gap is huge between these theories and designing a HRI

Social behavior and relationship are connected to perception, regulation and expression of emotions

- Emotion for social interactions have to fulfill two roles:
  - **Influencing decision processes** : Emotions change our decision process, favouring certain types of behaviours. Emotions can also modify the decision process in a more subtle way, just by altering the usual way an action is performed (facial micro expressions, quicker moves, etc.).
  - **Showing a desired or non-completely desired state of mind** : According to Goffman, people tend to play a role during social interactions. This role must be maintained to avoid ruptures. It is then really important to try to show coherent emotional state

The need to sense!



The need to sense!

from **RAW** data to **MEANING**

REFINERS → extract **semantics** from raw data

AGGREGATORS → combine raw data to:
extract **semantics**
make information **reliable**

The need to **sense**!

from **MEANING** to **EVENT**

LINKS → create **semantics** between information

AGGREGATORS + LINKS → construct to:

build rich information: **events**

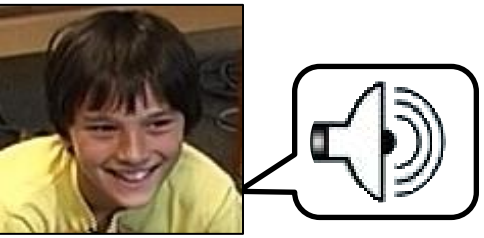
reason about information

Social and affective signals perception

Emotions play a central role in social interaction but **there are also several simultaneous levels of features :**

- **Long term features** are for example affective dispositions towards the others or personality features,
- **Medium term features** are more or less temporary states like the interactional signals such as role in interaction, positive / negative attitude, empathy,
- **Finally, short term features** correspond to the mode of interaction: emotion or mixture of emotion-related states such as stress, interest, confidence, uncertainty, deception, politeness, frustration, sarcasm, etc.

Multi-level detection of the emotional and interactional cues from speech



Low-level : acoustic cues

F0 variations

Energy variations

Rhythm

Duration

Speech/not speech

Emotion label
Pos./Neg. (*valence*)
Active/Not active
Average Valence
Average Activation

Dysfluences (*hesitations, silence, etc.*)
Talkative/Not
Reaction Time
Speech/not speech ratio

Mid-level : speech and emotion cues

High-level : user profile (Mind theory)

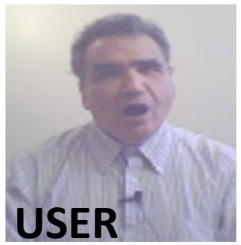
Emotional representation

Extraversion, Optimism, Emotionality, Self-Confidence

Interactional representation

Affinity, Dominance

Full-system – FUI ROMEO project



PERCEPTION of the robot:
Paralinguistic cues

Low-level clues (MFCC, prosodie, etc.)

Perception (Emotions, Speaker ID, Emotional, personality and interactional clues)

ROBOT

EMOTIONAL system (spirops AI)

User model

Internal emotion

Emotion
Personality
Interaction

Memory

Behavioural module

Dialog strategy

MOTORS

Verbal and multimodal actions

LIMSI team in Affective and social dimensions of spoken interactions

L. Devillers [1,2], A. Delaborde [1,3], M. Tahon [1,3], M. Soury [1,3] C. Barras [1,3]

LIMSI-CNRS (France) [1] / Univ. Paris-Sorbonne 4 [2] / Univ. Paris-Sud Orsay [3]

VIDEO : "Vivre avec les robots" – @ Elodie Fertil

How to use the CPM to understand affective states using the perception of audio and multimodal signals ?

The component process model (CPM) proposed by Scherer suggests 4 major appraisal objectives (Events include actor, target and action essential to compute emotion) :

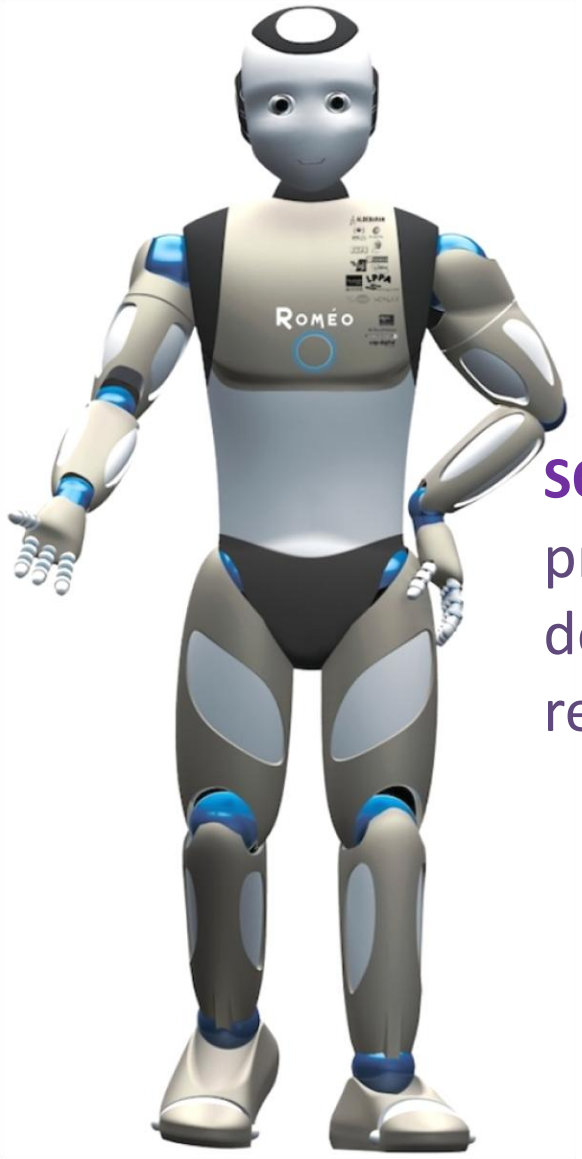
- Relevance : How relevant is this event for me? Does it directly affect me or my social reference group?
- Implications : what are the implications or consequences of this event and how do they affect my well-being and my immediate or long-term goals?
- Coping potential: how well can I cope with or adjust to these consequences?
- Normative significance: what is the significance of this event for my self-concept and for social norms and values?

Not straightforward to link perception and CPM without understanding semantic content?

- How to link non verbal signal perception such as voice quality cues and CPM :
  - tension or phonatory effort may correspond to sympathetic arousal
  - phonation perturbation and phonatory frequency may represent the “ability to control” part of the coping potential dimension.

How the CPM synchronizes the appraisal of a new feeling and an old one or mixtures of emotions ?

Guess intentions



intentions = goals behind

PRIMARY GOALS

often common to people

use a reflexive mechanism

SOCIAL GOALS

predefined by existing social relations

defined by the current social interaction nature

refined by the point of view of each participant

COMPLEX MECHANISMS

anticipate the effect of the speech act

try to deduct the purchased goal

Social decision

Goffman frames

=

Structured interaction

SOCIAL ROLES

goals, expected behaviors

INTERACTIONS

completion, mutation, …

REDUCE THE COMPLEXITY!

Social decision: emotions

EMOTIONS INFLUENCE DECISIONS

emotions trigger actions

EMOTIONS INFLUENCE PERFORMANCE

emotions change the way to perform actions

DECISIONS INFLUENCE EMOTIONS

show the right emotions to match primary and or social goals

EMOTIONS OVERRIDE DECISIONS

unintentional moves

Social decision: special behaviors

LIES

butler lie guided by social rules
to better understand lies

COMPASSION

evaluate emotions ➔ share emotions
strong affect binding

JOKE

repertoire with context (and mood)
good socializer

MECHANISM

tricks may be efficient
model of other's knowledge

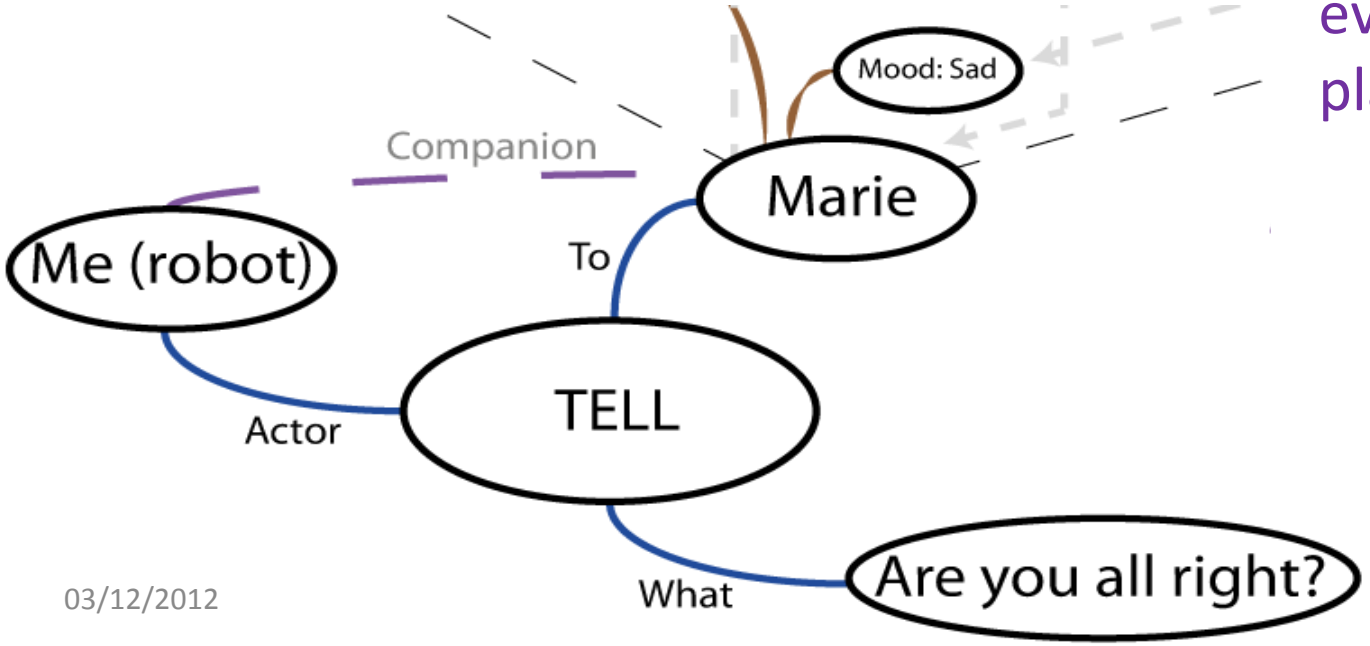
memories

common memories ➔ affective bond

Reminder as primary service

Sharing events

Granularity
From raw data to events or even plans



Network
concepts
attributes
time
links

memories management

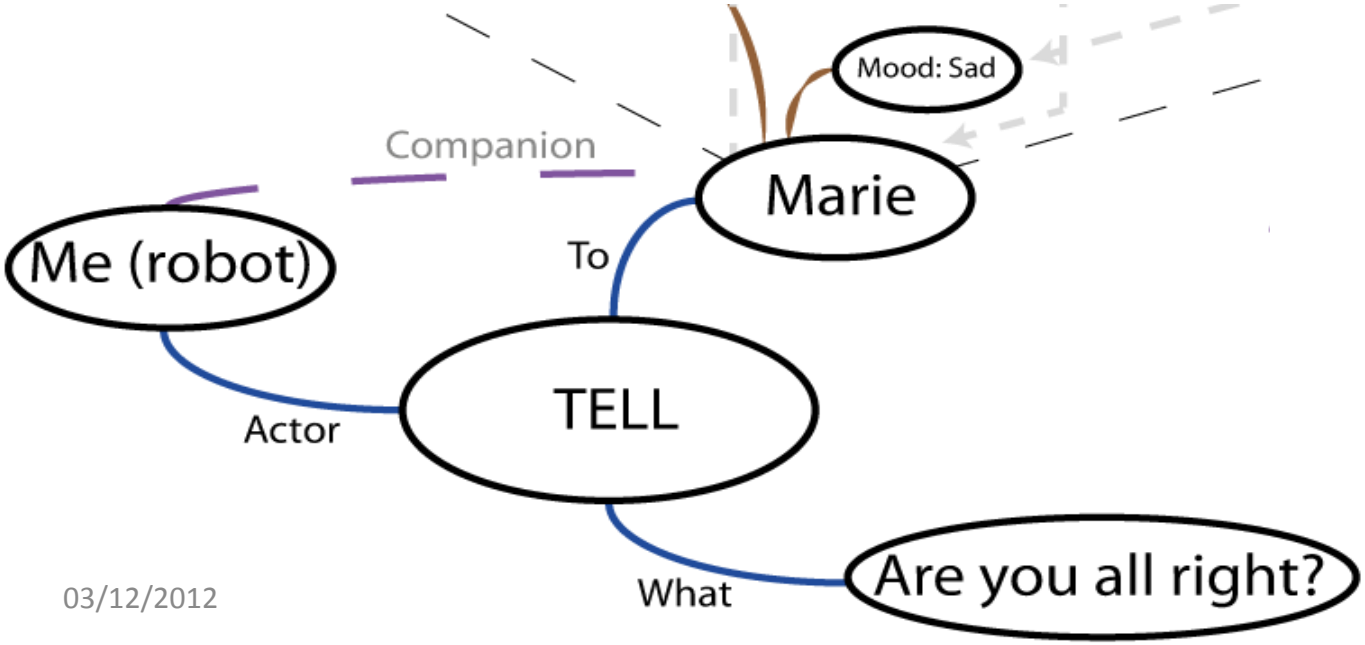
evaluate memories to filter & forget

built-in
security

emotional potential
(Scherer)

goals
(Scherer)

dependencies
important node
connected
usage



memories management

forget memories

merge

minimal information loss

forget

important information loss

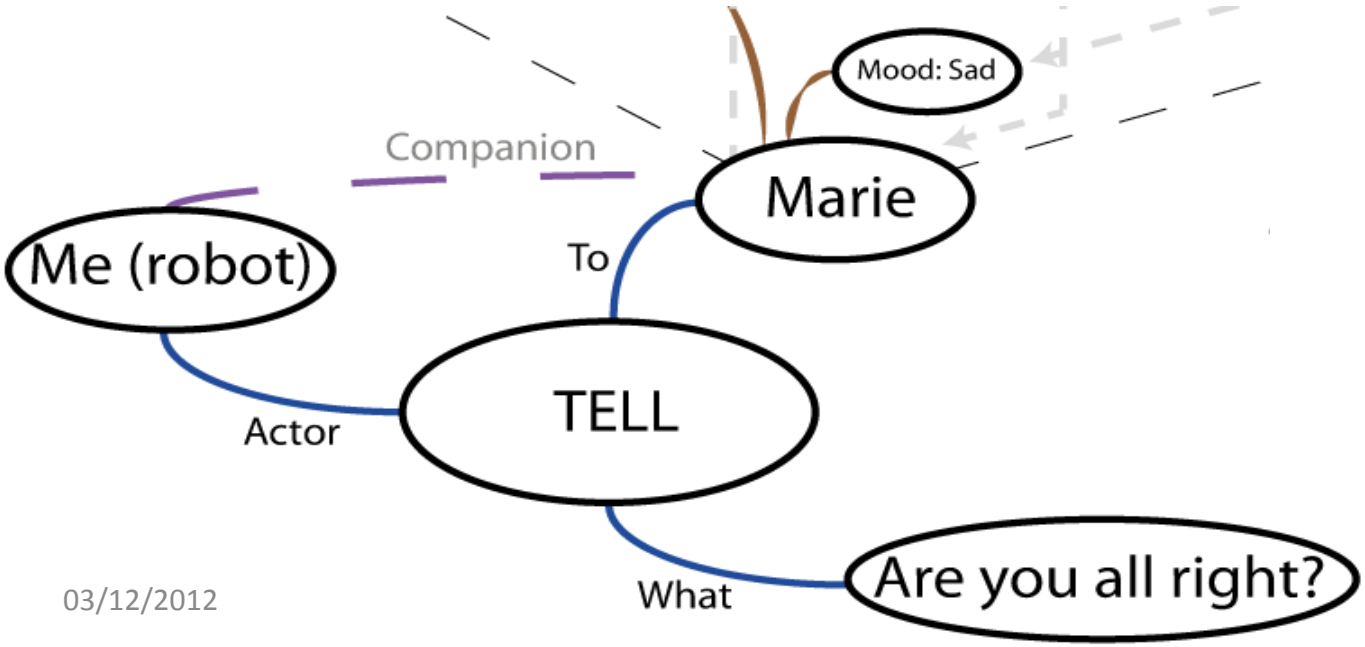
merge operators

time

periodicity

generalization

delete



Future trends & Challenges

Challenges:

- Social intelligence for HRI : perception, reasoning and generation of non verbal and verbal communication for real applications with different people: elderly, disabled people, etc.
- Need for new theories, knowledge (ex: links emotion/cognition), ontologies (emotion/type of tasks), memories but also empirical data and evaluation measures

We proposed some original ideas inspired from different theories for adding new capacities :

- a bio-inspired architecture (Mountcastle) : loop between perception / decision / anticipation
- Emotion and social traits extracted from signal (acoustic and multimodal) can be used for anticipation and memorization (mind theory, CPM Scherer)
- Social concepts such as roles/social goals (Goffmann) are used to simplify the anticipatory mechanisms

Future projects

- **The future of robots** in our society is certainly not to be seen as a replacement of human beings but as a new tool to simulate memory, educational assistance and mediation processes.
- Human beings behaviors such as **lies, compassion and jokes** imply that the robot has the ability to represent and understand some complex human beings behaviors
- **Emotions play a central role**: They are used to ease social interactions, to manage the short, mid and long term state of the robot, giving interlocutors the illusion of life, not only for a brief demo, but for a long term relationship.