

# Resolution of Conflicting Behaviors in Multimodal Believable Agents



Catherine Pelachaud

## Our goal

- Building a new human-interface for Information Delivering applications
- Dialoging with a pretty face is not enough
- Agent needs to be endowed with human-like qualities
  - Convey complex messages
  - Express messages verbally and nonverbally

## Believable Agent

- Definition Believable agent:
  - agent that acts consistently with her goals, state of mind, affect and personality.
  - Agent's behavior consistent with the meaning of her says
  - Agent's behavior synchronized with her speech
- Requirements:
  - Discourse: reflection of her state of mind
  - Expression: display or not display of an expression based on the social context

## BDI Agent

- BDI agent: Belief Desire Intention
- Mental state of the agent includes a representation of the beliefs and goals that drive:
  - the feeling of emotions
  - the decision of whether to display or not an expression

# What is Social Context?

- **Domain:** application-specificity, the events, action, objects in the world and their emotional effects on the Agent, ...
- **Cognitive Context:** information relevant to decision
  - Agent's display motive
  - Agent and Interlocutor's model
  - Agent and Interlocutor role and personal relationship
  - Social Interaction Context
- **Physical Context:**
  - Body
  - Device
  - ...

# Multimodal Signals

- Definition of a communicative act:
  - Action performed by Speaker
  - Goal: Addressee gets some beliefs about Speaker's beliefs and goals
- Characterization of multimodal signals by their placement with respect to linguistic utterance and significance in transmitting information. Eg:
  - Raised eyebrow may signal surprise, emphasis, question mark, suggestion...
  - Smile may express happiness, be a polite greeting, be a backchannel signal...

## Nonverbal communicative acts

- Need two information to characterize multimodal signals:
  - Their meaning
  - Their visual action
- Signal side → Muscular contraction on 3D facial model
- Meaning side → Set of beliefs that Speaker has the goal of communicating

# Nonverbal Communicative Act (Isabella Poggi)

- Need to define a semantic topology of types of information to be conveyed
- Four broad classes of meaning:
  - Information about Speaker's belief
  - Information about Speaker's intention
  - Information about Speaker's affective state
  - Meta-cognitive information about Speaker's mental action



## Information about Speaker's Belief

- Speaker has the goal that Addressee believes some specific speaker's beliefs
- This class encompasses several functions
  - Certainty markers: indicate how reliable the information is provided
  - Belief relation markers: denote relation between beliefs
  - Adjectival markers: mimic object dimension

## Information about Speaker's Intention

- Information about intention of single communicative act: performative of a sentence
- Information about a whole hierarchy of intentions: topic/comment
- Information about overall arrangement of conversation: turn-taking
- Direct Addressee's attention to events

## Information about Speaker's Affective State

- Emotions are **triggered** by event, action, person
- One can feel emotion toward another person (love, scorn)
- Emotion triggered by an event and not directed toward someone (fear, surprise)
- Signal: facial expressions of emotion
- Display of emotion are **regulated**
- Display or not of the emotion based on context

## Metacognitive Information about Speaker's Mental State

- Meta information about the source or cognitive states of information the speaker is taking about: gaze of thought.
- Try to make inferences about some fact  $p$
- Try to remember some fact  $p$

## Expression meaning

- **deictic**: this, that, here, there
- **adjectival**: small, subtle, difficult,...; or big, long, great,...
- **certainty**: certain, uncertain...
- **metacognitive**: I'm planning, I'm thinking, I'm remembering
- **performative**: greet, request (implore, order, suggest...),  
inform (warn, approve...),
- **topic comment**: this is the topic, this is the comment
- **Belief relation**: contrast,...
- **turn allocation**: take turn, give turn
- **affective**: anger, fear, happy-for, sorry-for, envy, relief, ....

## Expression signals

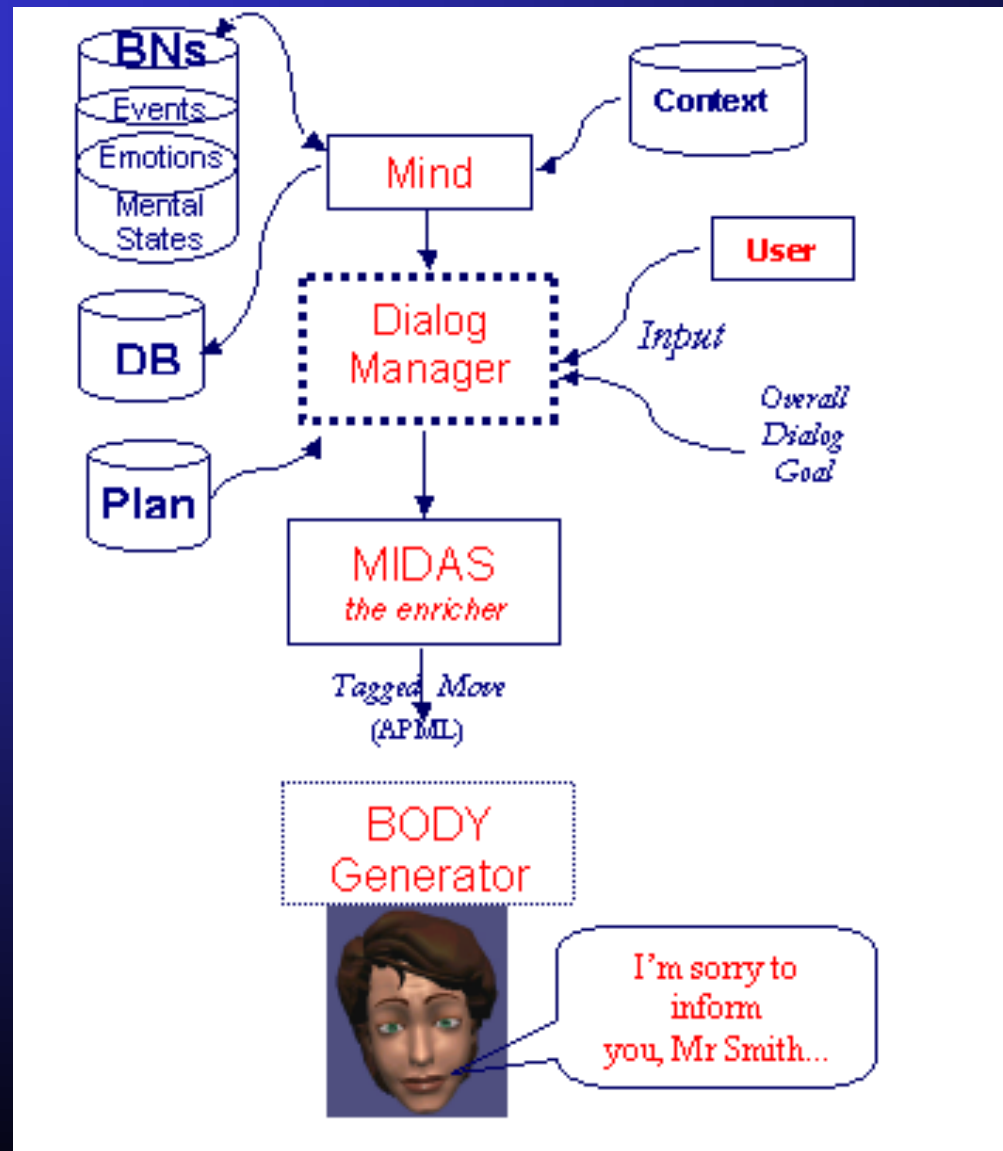
- **Deictic:** gaze direction, head direction
- **Certainty:**
  - Certain: small intensity frown
  - Uncertain: raised eyebrow
- **adjectival:**
  - Small, tiny...: small eye aperture
- **Belief relation:** contrast: raised eyebrow
- **Performative:**
  - Suggest: small raised eyebrow, head aside
- **Emotion:**
  - Sorry-for: head aside, inner eyebrow up
- ...

# System Overview

Fiorella de Rosis – Berardina De Carolis

- Information-delivering dialogs between Agent / User
- Process
  - Specify overall dialog goal
  - Select first communicative goal and say it
  - Give the speaking turn of User
  - User move is translated into symbolic communicative act
  - Decide what to say next
  - Cycle 2 to 4

# System Architecture





# MIND

- Update Agent's mental state:
  - Decide whether an affective state is activated
  - Decide if affect is shown based on social context
- Based on Dynamic Belief Network that combines:
  - BN of Agent's mental state at previous dialog turn
  - BN of Agent's mental state at the current dialog turn
  - BN that monitors the events during time interval and their possible causes and effects
- Return if necessary an affect associated to a CA

# MIND

- DBN of Agent's mental state
- 3 types of nodes
  - Belief
  - Goal
  - Goal achievement (with a weight)
- Personality modeled by different weights on different nodes
- Emotions caused by the belief that an important goal will fail or will be achieved.

# DIALOG MANAGER - DM

- Controls dialog flow
- Takes as input: overall dialog goal, i.e. main topic of discussion
- Decomposition in a set of communicative goals
- This set is transformed in a discourse plan using:
  - KB on the domain
  - Library of recipes: decomposition of communicative goal in sub-goals using a tree structure
  - Leaves: performative (ask, request, inform, greet...)

# MIDAS

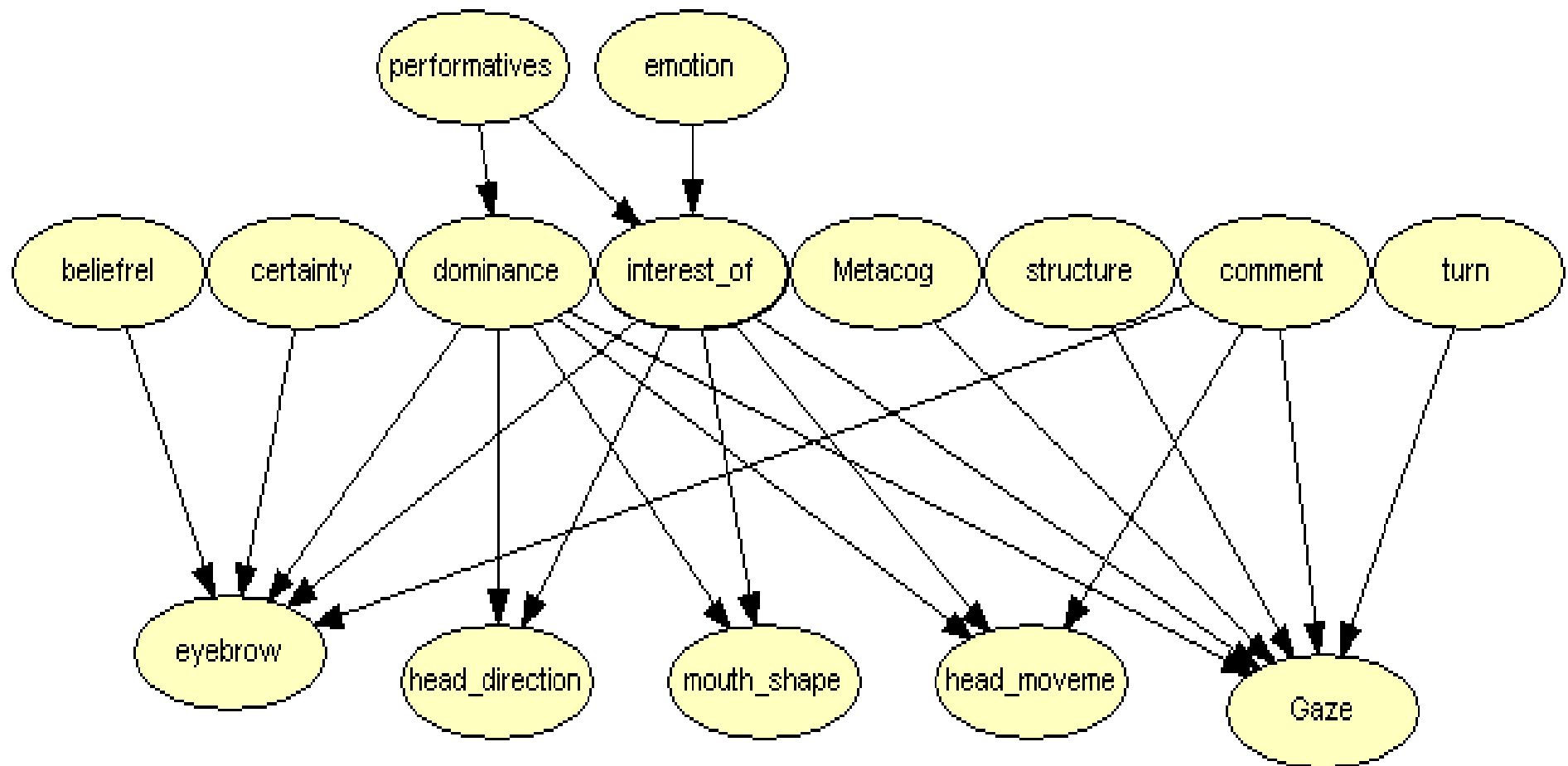
- Discourse plan is enriched with nonverbal communicative function
- Use the DBN built in Mind and the social context to decide whether to display an expression or not
- Insert all other tags using the information from discourse plan: Rethorical Relation, Focus, Performative...
- Output an XML tagged file using APML (Affective Presentation Markup Language) specification

## Body Generator

- Interprets APML tagged dialogs, i.e. all communicative functions
- Decides which signals to convey on which channels for each Communicative Act
- Synchronizes the signals with speech

## Conflict Resolution

- Different CAs may be shown on same facial channel with different values
- Belief Network links CAs to facial channels
- 2 types of nodes:
  - Root nodes: CAs
  - Leave node: signals
- 2 intermediate nodes: linking performative and emotion to signals
  - Dominance
  - Interest of



- Output of the belief network: value for facial channel
- Co-occurring expressions mix with each other at the facial channel level

<APML> <performative type = "inform" affect = "sorry-for" certainty="certain">

I'm sorry to tell you that you have been diagnosed as suffering from what we call angina pectoris,

</performative>

<belief-relation type = "elaboration-object-attribute">

which

<performative type = "inform" certainty="certain">

appears to be

<adjectival type = "small">

mild.

</adjectival> </performative> </belief-relation> </APML>



# Conflict Resolved

- *Sorry-for* is expressed by:
  - Head aside
  - Inner raised eyebrow central (eyebrow of sadness)
- *Certain* is expressed by:
  - Frown
- There is a conflict on the eyebrow channel between both signals of the Affective and Certainty Communicative acts
- The output of the BN is:
  - Head aside (*sorry-for*)
  - Frown (*certain*)



# Conflict Non-Resolved

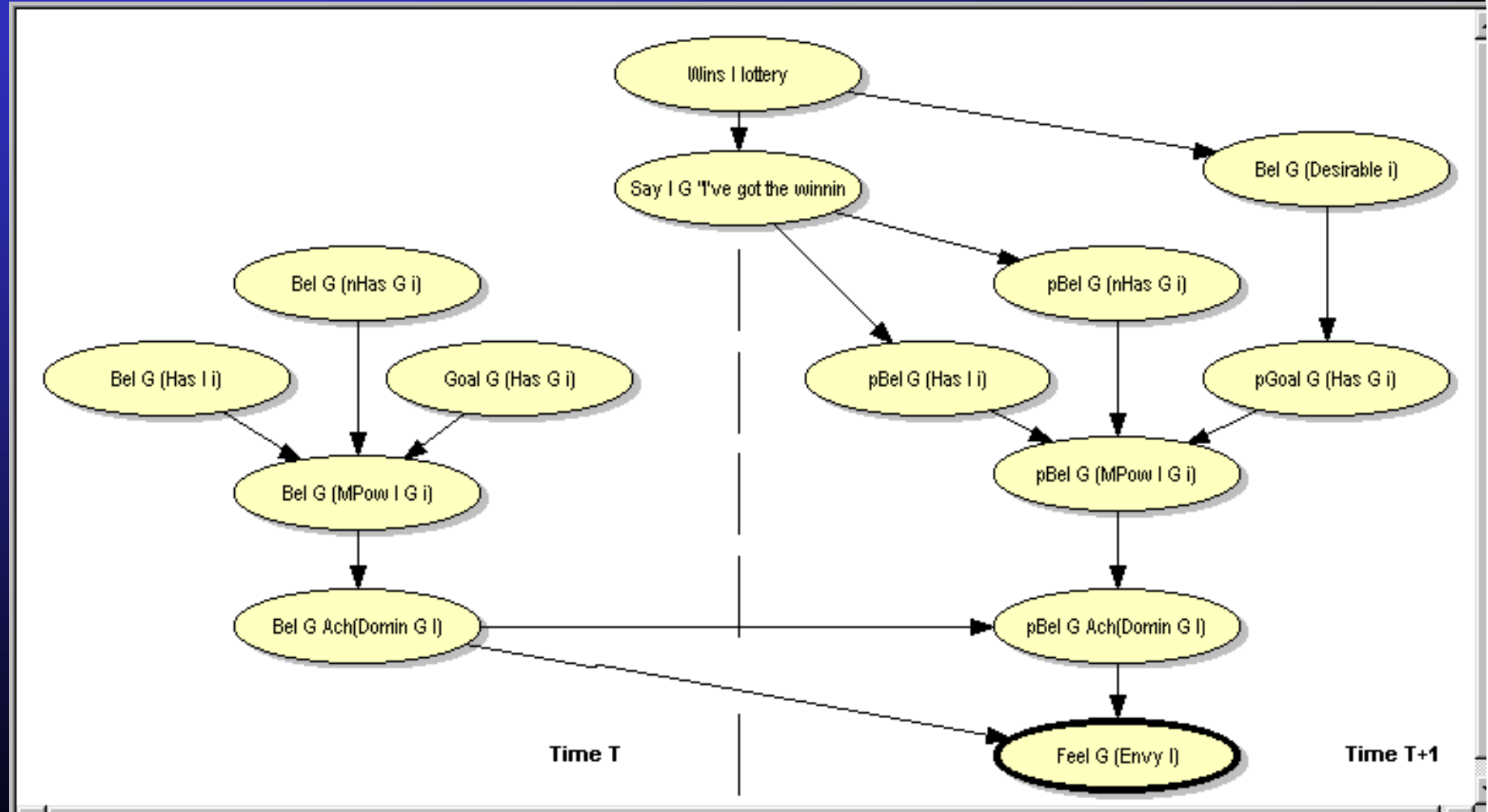
- Only the expression of *sorry-for* is shown



## Conclusion

- What are the minimum requirements to get a believable agent?
- What makes an agent coherent with her personality, social context...?
- What makes an agent an individual?
- How to embed gender and culture differences?
- How to make an agent reactive to user's nonverbal behavior (gaze, smile...)?

# BN for triggering ENVY



# Conclusion

- Definition of semantic topology of nonverbal behaviors
- Verbal and nonverbal discourse planner
- Notion of social context
- Notion of Conflict resolution

# Domain and Context Model

- Cognitive Context: information relevant to decision
  - Scenario Factors
  - Agent and Interlocutor' s models
  - Social Interaction Context and Practical Resources
- **Context**: information relevant to decision
  - Speaker's display motive: *vent, emphaty, consolation..*
  - Speaker's personality: *impulsive / non-impulsive*
  - Speaker - Addressee Role relationship:  
*neutral, S / A has power on A / S*
  - Speaker - Addressee Personal Relationship:  
*neutral, S / A aggressive A / S, S / A adoptive A / S*
  - Type of social interaction: *Public / intimate*
  - Addressee's model:  
*Cognitive capacity*  
*Personality*



# Cognitive Context

- **Agent's Display Motive**: the reason that induces the agent to display something in a particular situation (vent, empathy, advice, help, consolation, ...)
- **Agent's Personality**
- **Interlocutor's Features**
  - personality
  - cognitive capacity (comprehension, experience, problem solving)
- **Agent-Interlocutor Relationship**: Role - Personal
- **Type of Social Interaction**