Study Human Conversation
Study Human Conversation

• Goal
  • Understand and model “Embodied Conversation”

• Because we want to (1 of 3)
  • Create autonomous characters that can have natural, engaging and effective face-to-face conversations with humans (embodied conversational agents)
Study Human Conversation

- **Goal**
  - Understand and model “Embodied Conversation”

- **Because we want to (2 of 3)**
  - Create systems that allow remote people to communicate with each other in online environments as if they were meeting face-to-face (avatars)
Study Human Conversation

- **Goal**
  - Understand and model “Embodied Conversation”

- **Because we want to (3 of 3)**
  - Create and animate believable virtual actors and extras in scenes involving language interaction (character animation)
Building an Animation Tool: BEAT
Building a System: BEAT

• Approach

Takes as input the lines to be spoken by characters and a description of a scene, tries to understand the conversation functions being carried out and then produces supporting nonverbal behavior. The output is an animation script with gesture synchronized with speech.
(GA) It is apparently (GT) \{ some \[ \text{1 kind} \] of \} \{ \[ \text{2 virtual} \[ \text{3 actor} \] \] \}.

TAGGED

You just need to type in a line like "This is Bradley Beat live from SIGGRAPH", and the actor is able to talk and gesture by itself.
BEAT Architecture: Script Input

Script Feeder

Context
- Language Module
- Generation Module
- Filtering Module

Speech Timing
- Scheduling Module
- Animation Compiler

Generations: GEN 1, GEN 2, ..., GEN n
Filters: FIL 1, FIL 2, ..., FIL n

“This is both good news and bad news”
This is good news and bad news both.
Anyway,

The tile is blue

Anyway

The tile is blue

good

bad

the thin guy

walked away

this sucks Joe
BEAT Context Representation

- **Domain Knowledge**: World and gesture ontology
- **Scene Description**: What and who can I see
- **Participation Framework**: What is everyone’s role
- **Dictionary**: How are terms related
- **Discourse History**: What has been said
This is good news and bad news both.
BEAT Architecture: Behavior Generation

- Context
- Generation Module
- Filtering Module
- Scheduling Module
- Speech Timing
- Animation Compiler

Script Feeder
Language Module
Generation Module
- GEN 1
- GEN 2
- GEN n

- FIL 1
- FIL 2
- FIL n

CLAUSE

THEME
OBJECT
ACTION
NEW

RHEME
OBJECT
NEW
CON(1)
GEST: BEAT
NEWS
NEW
CON(1)
BAD
NEWS

This is good news and bad news both.
BEAT Architecture: Behavior Generation

- Context
- Generation Module
- Filtering Module
- Scheduling Module
- Animation Compiler
- Script Feeder
- Language Module
- Generation Module
- Filtering Module
- Scheduling Module
- Animation Compiler

TONE: BREAK

TONE: ENDHI GAZE: AWAY

TONE: ENDLO GAZE: HEARER

EYEBROWS: RAISED

ACCT: HI
HD: NOD
GEST: BEAT
GEST: CON_R

This is both good news and bad news...
BEAT Architecture: Behavior Filtering

This is good news and bad news both.
This is good news and bad news both
This is good news and bad news both
This is good news and bad news both...
<AnimationScript SPEAKER="AGENT" HEARER="USER">
<START SPEECH="This is both good news and bad news">
<START ACTION="GAZE" DIRECTION="AWAY_FROM_HEARER" SRT="0.0">
<START ACTION="VISEME" TYPE="B" SRT="0.0">
<STOP ACTION="GAZE" DIRECTION="AWAY_FROM_HEARER" SRT="0.511">
<START ACTION="VISEME" TYPE="A" SRT="0.511">
<START ACTION="GAZE" DIRECTION="TOWARDS_HEARER" SRT="0.801">
<START ACTION="EYEBROWS" SRT="0.801">
<START ACTION="GESTURE_RIGHT" TYPE="CONTRAST_1"
RIGHT_TRAJECTORY=CONTRAST_TRAJECTORY
RIGHT_HANDSHAPE="CONTRAST" SRT="0.801">
BEAT Example: Maya Compilation

Script Feeder → Language Module → Generation Module → Filtering Module → Scheduling Module → Maya MEL Compiler

Context

Speech Timing

GEN 1, GEN 2, ..., GEN n, FIL 1, FIL 2, ..., FIL n
BEAT Example: Animator Instructions

“\text{You just need } [1 \text{ to *type in } ] \{ \text{ a } [2 \text{ *line } ] \text{ like } \} ”

[1] ICONIC - Typing action

* Pitch accent
{} Raised eyebrows
. . Gaze away
___ Gaze towards