1. Introduction

-- Communicative Intelligence --

Toyoaki Nishida
Kyoto University
<table>
<thead>
<tr>
<th>Year</th>
<th>AI</th>
<th>ICT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>AI</td>
<td>ICT</td>
</tr>
<tr>
<td>------</td>
<td>-------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Theory</td>
</tr>
<tr>
<td>1950~</td>
<td>1952-62: Checker program by A.Samuel 1956: Dartmouth Conference</td>
<td>1957:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Universal Turing Machine - from Hardware to Software</td>
</tr>
<tr>
<td>1960~</td>
<td>Heuristics - Intelligence as Clever Computing</td>
<td>1961:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interdisciplinary Science - Artificial Intelligence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1963: Interactive Computer Graphics by I.Sutherland</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1968: Mouse and Bitmap display for on-line System (NLS) by D.C.Engelba</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1969: ARPA-net</td>
</tr>
<tr>
<td></td>
<td>Combinatorial Explosion problem pointed out in The Lighthill report</td>
<td>1972:</td>
</tr>
<tr>
<td></td>
<td>by D.Marr 1977: Automated Mathematician (AM) by D.Lenat 1978:</td>
<td>1979: Expanding Internet World</td>
</tr>
<tr>
<td></td>
<td>Autonomous Vehicle Stanford Car by H.Meger</td>
<td></td>
</tr>
<tr>
<td>1980~</td>
<td>Symbolic Computing - Knowledge-based Systems</td>
<td>1980:</td>
</tr>
<tr>
<td></td>
<td>Vehicle ALVINN by D.Pomerleau</td>
<td></td>
</tr>
<tr>
<td>1990~</td>
<td>1990: Genetic Programming by J.R.Koza</td>
<td>1992:</td>
</tr>
<tr>
<td></td>
<td>Embodied &amp; Network Intelligence - Like human and society</td>
<td>1994:</td>
</tr>
<tr>
<td></td>
<td>2000: Honda Asimo</td>
<td>1997:</td>
</tr>
<tr>
<td></td>
<td>2004: The Mars Exploration Rovers (Spirit &amp; Opportunity)</td>
<td>1998:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Generic Computing Engines - from Software to Data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1998: W3C (eXtensible Markup Language) by T.Berners-Lee</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1998: PayPal</td>
</tr>
<tr>
<td></td>
<td>2012: Siri</td>
<td>2003: Skype / iTunes store</td>
</tr>
<tr>
<td>2010~</td>
<td></td>
<td>2004: Facebook</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2005: YouTube / Google Earth</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2006: Twitter</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2007: Google Street View</td>
</tr>
</tbody>
</table>
Communicative Intelligence for Bridging People and SI

Super Intelligence

People

Communicative Intelligence

[Nishida-Nakazawa-Ohmoto-Mohammad 2014]
Long-term goal

Challenge: A robot that can participate in conversation
Conversation is a complex business

Trust
Conviviality
Social networks

Asking
Proposing
Negotiating

Eye gaze
Facial expression
Hand gesture
Posture
Conversational Informatics

Application

Content production

Platform

Model building

Evaluation

Measurement

Analysis

Theory

Conversational interactions

[Nishida-Nakazawa-Ohmoto-Mohammad 2014]
**Applied Intelligence Information Processing Group**

**Interaction Measurement, Analysis and Modeling**
We aim at uncovering principles of verbal and nonverbal interactions that people engage everyday as a part of intellectual activities and developing a suite of technologies to help people benefit from conversational interactions.

**Cognitive Design**
We aim at designing the expressions, functions, control, and interactions of artifacts based on a physio-cognitive approach.

**Augmented Conversation Environment**
We aim at building a smart environment that integrates conversation in physical and virtual spaces.

**ICIE: Immersive Collaborative Interaction Environment**
Conversation in the physical environment

**IMADE: Interaction Measurement, Analysis, and Design Environment**
Conversation in the virtual environment
ICIE for the shared virtual world

https://www.youtube.com/watch?v=V-9SKpcMrzk

[Lala 2013]
ICIE – immersive collaborative interaction environment

[Nishida-Nakazawa-Ohmoto-Mohammad 2014]
Immersive WOZ environment

Feedback generation
Motion mapping
User motion sensing
Head recognition
Gesture recognition
Face model
Human body model

Tele-operated robot

The conversation place

WOZ operating environment

[Ohmoto 2013a]
Projecting the real world into the virtual world

http://www.youtube.com/watch?v=68UrJv65HvY

[Ohmoto 2011]
IMADE: Interaction Measurement, Analysis, and Design Environment
iCorpusStudio

List of analysis target

Annotations

[Nishida-Nakazawa-Ohmoto-Mohammad 2014]
Collaborative annotation system

[Annotation]

Mutually sharing the subjective criteria to preserve consistency.

Share subjective annotation rules

Construction of annotation data structure

Feature points of annotations are statistically extracted.

Construction of annotation rule structure

Provide

Submit

[Annotation]

[Nishida-Nakazawa-Ohmoto-Mohammad 2014]
3D conversation capture – over the shoulder
Non-intrusive Eye Gaze Tracking (EGT) by corneal imaging

Corneal Image Feature Matching

Problem:
Local feature correspondence + RANSAC does not work due to large noise in eye images

Approach:
1. Formulate problem as registration of 3D spherical light maps of eye and scene image
2. Single point algorithm for robust alignment

Corneal Imaging Camera

- Lightweight and versatile system
- Appl.: Google Glass like HMD, unconstrained setups

Scene camera
Eye camera

Scene Image
Eye Image

Spherical light map of a scene
Spherical light map of an eye
Application 1: Non-intrusive and uncalibrated PoG estimation

Scene images with PoG

↑ Aligned results (from eye images)

Eye images with GRP

Gaze trajectory in static scene image

Application 2: Peripheral vision estimation

Peripheral vision map in eye image

Peripheral vision map overlaid to scene image
Quantitative analysis of pupil dilations against the difficulty of the task

Estimate how Pupil dilations is related to the difficulty of the irritating maze game.

[Tange, Nakazawa 2015]
Pupil dilations in the irritating maze game

[Tange, Nakazawa 2015]
Pupil dilations in the irritating maze game

- Delay before contraction
- Expand before the narrow conduit

Pupil size (pixel) vs. Elapse time (ms)

[Tange, Nakazawa 2015]
Tacit social signals

Social signals
Facilitative actions

Intentions are dynamic and tacit

Facilitative agent that can track dynamic and tacit intentions

The first half

The second half

[Nishida-Nakazawa-Ohmoto-Mohammad 2014]
Agent Control

Facilitative agent

- Estimation of emphasizing points
  - Evaluating estimation process
  - Switch or not between divergent and convergent
    - Results
    - Social signals
    - Agent behavior
    - Select next proposal
    - Next proposal

Results

Estimation agent

- Estimation of emphasizing points
  - Results
  - Divergent
    - Social signals
    - Select next proposal
  - Agent behavior
  - Next proposal

[Ohmoto 2013b]
Facilitative agent is significantly more preferred in all aspects we have investigated.

[Ohmoto 2013b]
Inducing intentional stance toward agent players

Q: How can we induce an intentional stance toward NPCs?

H: Presenting goal-oriented trial-and-error process with multimodal behaviors.

[Ohmoto 2015]
Inducing intentional stance toward agent players

Video analysis

Trial-and-error agent:
Presents goal-oriented trial-and-error process by multimodal behavior (e.g., hand gesture, body orientation, etc that only ambiguously display mental states).

Text display agent:
Presents agent’s behavioral intention by unambiguous text.

Questionnaire analysis

[Ohmoto 2015]
Inducing intentional stance toward agent players

Q: How can we induce an intentional stance toward NPCs?
H: Demonstrating strategy change.

[Suyama, Ohmoto 2015]
Estimating user’s concentration during exercise

**Purpose**
Evaluating the physiological sensing method for estimating user’s concentration and appropriate advice giving during exercise.

**Method**
VR exercise game as the task.
Using SCR and LF/HF.
Controlling the condition of advice giving.

**Findings**

- **SCR**
  - Hide from the chaser
  - Not concentrated

- **LF/HF**
  - Hide in the place the chaser checked previously.
  - Simply moving around

- **Deliberating but not reacting**
  - Advice

- **Deliberating and reacting**
  - No advice
Learning by imitation

1. Watch
   - Operator
   - Commands
   - Feedback
   - Actions
   - Learner
   - Model

2. Learn
   - Command stream
   - Action stream
   - Discovery Phase
     - Discrete Commands
     - Discrete Actions
   - Association Phase
   - Behavior Generation Model
   - Controller Generation
   - Robot/Agent Controller

3. Act
   - Operator
   - Commands
   - Feedback
   - Actions
   - Learner
   - Induced Model

[Mohammad 2009]
Constrained Motif Discovery:
- Given a time series $X(t)$
  find recurring patterns of length between $L_1$ and $L_2$ using distance function $D$ subject to the constraint $P(t)$, where $P(t)$ is an estimation of the probability that a motif occurrence exists near time step $t$. 

[Mohammad 2009]
Learning by imitation

Robust Singular Spectrum Transform

\[ H(t) = \left[ \text{seq}(t-n); \ldots; \text{seq}(t-1) \right] \]

\[ G(t) = \left[ \text{seq}(t+1); \ldots; \text{seq}(t+n) \right] \]

\[ H(t) = U(t)S(t)V(t)^T \]

Find optimal \( l_P \)

\[ G(t)G(t)^T u^g = \mu u^g \]

Find optimal \( l_F \)

\[ \beta_i(t) = u^g_i, \quad i \leq l_F \quad \text{and} \quad \lambda_j - 1 \leq \lambda_j \leq \lambda_{j+1} \]

\[ \alpha_i(t) = \frac{U_iU_i^T \beta(t)}{\|U_iU_i^T \beta(t)\|}, \quad i \leq l_f \]

\[ cs_i(t) = 1 - \alpha_i(t)^T \beta_i(t) \]

\[ \hat{x}(t) = \sum_{i=1}^{l_f} \lambda_i \times cs_i \]

\[ \tilde{x}(t) = \hat{x}(t) \times \left| \mu_F(t) - \mu_P(t) \right| \times \left| \sqrt{\sigma_F(t)} - \sqrt{\sigma_P(t)} \right| \]

[ Mohammad 2009]
Learning by imitation

Learning Interaction Protocols using Augmented Bayesian Networks Applied to Guided Navigation

Yasser Mohammad and Toyoaki Nishida

IEEE IROS 2010
Effect of Mutual Imitation on Perception of Robot's Imitative Skills

Yasser Mohammad
Toyoaki Nishida
Synthetic evidential study (SES) combines dramatic role play and group discussion to help people spin stories by bringing together partial thoughts and evidence.
At the beginning of the 18th century, a feudal lord named Asano Takumi-no-kami Naganori was in charge of a reception for envoys from the Imperial Court in Kyoto. Another feudal lord, Kira Kozuke-no-suке Yoshinaka, was appointed to instruct Asano in the ceremonies. On the day of the reception, while Kira was talking with Yoriteru Kajikawa, a lesser official, at “Matsu no Roka” (“Hallway of Pine Trees”) in Edo Castle, Asano came up to them screaming “This is for revenge!!” and slashed Kira twice with a short sword. Soon after the incident, Kajikawa restrained Asano, who was then imprisoned. The reason for the attack was not known, though it was widely believed that Kira had somehow humiliated Asano. Ultimately Asano was sentenced to commit seppuku, a ritual suicide, but Kira went without punishment.
Dramatic Role Play
Group play capture
Agent play
Discussion phase

Third person view

First person view

Asano

Kira

Kajikawa

Discussions
History of conversational systems development

- Natural language dialogue systems
- Story Understanding systems
- Natural language question answering systems
- Speech dialogue systems
- Multi-modal dialogue systems
- Embodied Conversational Agents / Intelligent Virtual Human
- The Knowledge Navigator
- Interactional systems
- Conversational Systems
- Transactional systems
- Affective Computing
- Cognitive systems

Timeline:
- 1970
- 1980
- 1990
- 2000
- 2010
Early Natural Language Dialogue Systems

**Baseball**

- One of the earliest natural language question answering system.
- Answers such questions about baseball games.

**Specification list:**

“Where did the Red Sox play on July 7?”
- Place = ?
  Team = Red Sox
  Month = July
  Day = 7

“What teams won 10 games in July?”
- Team(winning) = ?
  Game(number of) = 10
  Month = July

**Dictionary:**

“team” -> meaning: Team=(blank)
“Red Sox” -> meaning: Team=Red Sox
“who” -> meaning: Team=?
“winning” -> meaning: subroutine

**Data:**

Month = July
Place = Boston
Day = 7
Game Serial No. = 96
(Team = Red Sox, Score = 5)
(Team = Yankees, Score = 3)

[Green 1961]
Early Natural Language Dialogue Systems

Modules of Baseball

1. Question Read-in
2. Dictionary Look-up
3. Syntax
   (1) scan for ambiguities in part of speech: in some cases, resolved by looking at adjoining words; in other cases, resolved by inspecting the entire question.
   (2) locates and brackets the noun phrases, [], and the prepositional and adverbial phrases, ().
   The verb is left unbracketed. E.g.,
   “How many games did the Yankees play in July?”
   -> [How many games] did [the Yankees] play (in [July])?

Any unbracketed preposition is attached to the first noun phrase in the sentence, and prepositional brackets added. E.g.,

“Who did the Red Sox lose to on July 5?”
-> (To [who]) did [the Red Sox] lose (on [July 5])?

The syntax routine determines whether the verb is active or passive and locates its subject and object.

The syntax analysis checks to see if any of the words is marked as a question word. If not, a signal is set to indicate that the question requires a yes/no answer.

4. Content Analysis
   The content analysis uses the dictionary meanings and the results of the syntactic analysis to set up a specification list for the processing program. E.g.,

   “each team”: Team=(blank) -> Team=each
   “what team”: Team=(blank) -> Team=
   “Who beat the Yankees on July 4”: Team=(blank) -> Team=?, Team(winning)=? Team(losing)=Yankees
   “six games”: Game=(blank)->Game(number of)=6
   “how many games”: Game=(blank)->Game(number of)=?
   “Who was the winning team...”?:
   Team=? and Team(winning)=(blank)
   -> Team(winning)=?

5. Processing
   The specification list indicates to the processor what part of the stored data is relevant for answering the input question. The processor extracts the matching information from the data and procedures, for the responder, the answer to the question in the form of a list structure.

[Weizenbaum 1966]
Early Natural Language Dialogue Systems

LUNAR

Prototype natural language question answering system that helps lunar geologists access chemical analysis data on lunar rock and soil composition.

(i) Syntactic analysis using heuristic/semantic information to choose the most likely parsing (Augmented Transition Network Grammar was used)
(ii) Semantic interpretation: produce a formal representation for queries
(iii) Execution of this formal expression in the retrieval
Early Natural Language Dialogue Systems

Augmented Transition Network Grammar

S:

[Diagram of S showing NP, V, AUX, and NP nodes with arrows indicating transitions]

NP:

[Diagram of NP showing DET, ADJ, NOUN, PRON, and DET nodes with arrows]

PP:

[Diagram of PP showing PREP, NP, and PP nodes with arrows]

[Woods 1973]
SHRDLU

- Natural language understanding system working on the “Blocks” world.
- Based on the belief that “a computer cannot deal with language unless it can understand the subject it is discussing”
- Answers questions, executes commands, and accepts information in normal English dialog.
- Uses semantic information and context to understand discourse and disambiguate sentences.
- Procedural knowledge representation <-> declarative KR

Early Natural Language Dialogue Systems

Early Natural Language Dialogue Systems

**ELIZA**: Natural language dialogue system without explicit knowledge about the discourse domain

Transformation rules: collection of key lists.

\[
(K \ ( (D_1) \ (R_{1,1}) \ (R_{1,2}) \cdots \ (R_{1,m_1}) ) \\
( (D_2) \ (R_{2,1}) \ (R_{2,2}) \cdots \ (R_{2,m_2}) ) \\
\vdots \\
( (D_n) \ (R_{n,1}) \ (R_{n,2}) \cdots \ (R_{n,m_n}) ))
\]

where, \( K \): keyword, \( D_i \): decomposition template, \( R_{i,j} \): Reassembly rule.

Example of a decomposition template and a reassembly rule:

\[
D: \quad (0 \ YOU \ 0 \ ME) \\
R: \quad (WHAT \ MAKES \ YOU \ THINK \ I \ 3 \ YOU)
\]

“It seems that you hate me”

-> “What makes you think I hate you”

Substitution rule. E.g.,

\[
(YOURSELF = MYSELF) \\
(MY \ YOUR \ 5 \ (\text{transformation rules} ))
\]

[Weizenbaum 1966]
The Hearsay-II Speech-Understanding System

Integrates multiple levels of information processing by knowledge sources coordinated by the blackboard model:

- Parameter
- Segment
- Syllable
- Word
- Word-sequence
- Phrase
- Data base interface

Combining top-down (hypothesis driven) and bottom-up (data-driven) processing.

Selective attention to allocate limited computing resources.

[Erman 1980]
The blackboard architecture

(a) Parameters

(b) Segments

(c) Syllable classes

(d) Words – 2

(e) Words – 1

(f) Word sequences

(g) Phrases

[Erman 1980]
The blackboard architecture

Level_{k+1} \rightarrow \ldots \rightarrow Level_2 \rightarrow Level_1

Blackboard

Focus-of-control

Database

Scheduler

Scheduling queues

Blackboard Monitor

KS_1 \rightarrow \ldots \rightarrow KS_n

Program modules

Databases

Data flow

Control flow

[Erman 1980]
Early Speech Dialogue Systems

PUT-TTHAT-THERE

The conjoint use of voice-input and gesture recognition to command events on a large format raster-scan graphic display in a “Media Room”

Commands

“Create a blue square there.”
“Move the blue triangle to the right of the green square”
“Move that to the right of the green square.” (with pointing)
“Put that there” (indicated by gesture)
“Make that smaller” (with pointing gesture)
“Make that (indicating some item) like that (indicating some other item)”
“Delete that” (pointing to some item)
“Call that ... the calendar” (with pointing)

[Bolt 1980]
“Phil”: concrete image of a personal assistant.

[Apple Computer 1987]
Social responses to computers

- not: conscious beliefs that computers are human or human-like.
- not: from user’s ignorance or psychological or social dysfunctions
- not from: a belief that subjects are interacting with programmers
- rather: human-computer relationship is fundamentally social.

Nass 1994]
Requirements

- **Support interactive give and take**
  Assistants will need not only respond to questions but also ask questions.

- **Recognize the costs of interaction and delay**
  It is inappropriate to require the user’s confirmation of every decision made while carrying out a task.

- **Manage interruptions effectively**
  When it is necessary to initiate an interaction with the user, the assistant needs to do so carefully, recognizing the likelihood that the user is already occupied to some degree.

- **Acknowledge the social and emotional aspects of interaction**
  To become a comfortable working partner, a computer assistant will need to vary its behavior depending on such the task, the time of day, and the boss’s mood.

=> The PERSONA project @ Microsoft Research [Ball 1997]
The architecture of Peedy

- Spoken Language Interpretation
- Dialogue Management
- Character Animation
- Background and Object Animation
- Sound Generation

[Ball 1997]
The architecture of Peedy

Microphone
- Whisper
  - Speech Recognition

Character
- Player/ReActor
  - Animation Engine

Sound
- Speech Controller

Names
- Proper Name Substitution

NLP
- Language Analysis

Semantic
- Template Matching & Object Descriptions

Names Database

Action Templates Database

Object Database (CDs)

Speech & Animation Database

Dialogue
- Context & Conversation State

Dialogue Rules Database

CD Player

CD Changes

Speech Controller

Application

[Ball 1997]
A state transition machine model with 5 conversational states and 17 input events is used.

Peedy: “I have The Bonnie Raitt Collection, would you like to hear something from that?”

User: “Sure”

pePickTrack; genTracks; doSelect; Say <!PDSays(how About)>

pePickTrack: causes Peedy to look down at the CD (note) that he’s holding as if considering a choice.
genTracks: expands the description of the current CD into a list of the songs it contains.
doSelect: select one or two tracks, based on the parameters given in the interaction.
Say: verbally offer the selected song with the appropriate beak-sync.

[Ball 1997]
Believability

Believable agents

provides the illusion of life, and permits the audience’s suspension of disbelief.
... critical in theater, film, animation, radio drama, ...

Questions

i. What makes characters believable? What makes agents believable?
   Are agents different from characters?

ii. What is the nature of current integrated architectures and situated agents?
    How well do they support believability? Are there clear areas demanding study?

iii. How much breadth of capability is necessary to produce a believable agent?
    How much depth (competence) is necessary?

iv. What are the respective role of movement and language in achieving believability?

v. What is the role of context in establishing expectations in the user and thus is simplifying the task?
   For instance, how can the setting of an interface agent or the theme of a story help limit the technical requirements on agents?

vi. How can we measure believability and progress toward believability?
    Why don’t artists use “scientifically valid” techniques to evaluate the believability of their characters?

[Bates 1994]
Interactive Drama:
Interactive drama concerns itself with building dramatically interesting virtual worlds inhabited by computer-controlled characters, within which the user (the player) experiences a story from a first person perspective. ... (Bates 1992)

Drama = Character + Story + Presentation

[Mateas 1997]
Multi-modal dialogue engine

Graphics

Sound

Text

Cloud

Jennifer James

[Hayes-Roth 1998]
Oz project

**Personality**
Rich personality should infuse everything that a character does. What makes characters interesting are their unique ways doing things.

**Emotion**
Characters exhibit their own emotions and respond to the emotions of others in personality-specific ways.

**Self-motivation**
Characters have their own internal drives and desires which they pursue whether or not others are interacting with them.

**Change**
Characters grow and change with time, in a manner consistent with their personality.

**Social relationships**
Characters engage in interactions with others in a manner consistent with their relationship. In turn, these relationships change as a result of the interaction.

**Illusion of life**
Pursuing multiple, simultaneous goals and actions, having broad capabilities, and reacting quickly to stimuli in the environment.

[Mateas 1997]
Implements the social, linguistic, and psychological conventions of conversation.

- Has a human-like body. Uses eye gaze, body posture, hand gestures, and facial displays to organize and regulate the conversation.
- The conversational model relies on the function of non-verbal behaviors as well as speech.
- A full symmetry between input and output modalities: not only respond to visual, audio and speech cues (such as speech, shifts in gaze, gesture, and non-speech audio) but also generate these cues.
Rea

[Cassell 1999]
BEAT (Behavior Expression Animation Toolkit)

Discourse Model

Knowledge Base

Word timing

Behavior Generation

Behavior Suggestion

Behavior Selection

Behavior Scheduling

Generator Set

Filter Set

Translator

Animation

Text Input

Language Tagging

[Cassell 2004]
"Mary socked John."
"Mary punched John."
"Mary hit John with her fist."
"John was socked by Mary."
"Marie a donne un coup de poing a Jean."
"Maria pego a Juan."

<table>
<thead>
<tr>
<th>Primitive</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATRANS</td>
<td>Transfer of an abstract relationship (i.e. give)</td>
</tr>
<tr>
<td>PTRANS</td>
<td>Transfer of the physical location of an object (e.g., go)</td>
</tr>
<tr>
<td>PROPEL</td>
<td>Application of physical force to an object (e.g. push)</td>
</tr>
<tr>
<td>MOVE</td>
<td>Movement of a body part by its owner (e.g. kick)</td>
</tr>
<tr>
<td>GRASP</td>
<td>Grasping of an object by an action (e.g. throw)</td>
</tr>
<tr>
<td>INGEST</td>
<td>Ingesting of an object by an animal (e.g. eat)</td>
</tr>
<tr>
<td>EXPEL</td>
<td>Expulsion of something from the body of an animal (e.g. cry)</td>
</tr>
<tr>
<td>MTRANS</td>
<td>Transfer of mental information (e.g. tell)</td>
</tr>
<tr>
<td>MBUILD</td>
<td>Building new information out of old (e.g. decide)</td>
</tr>
<tr>
<td>SPEAK</td>
<td>Producing of sounds (e.g. say)</td>
</tr>
<tr>
<td>ATTEND</td>
<td>Focusing of a sense organ toward a stimulus (e.g. listen)</td>
</tr>
</tbody>
</table>

Action: PROPEL
Actor: Mary
Object: Fist
From: Mary
To: John

[Schank 1975]
Story understanding and generation

Script-based understanding

SAM (Script Applier Mechanism)

FRUMP
Story understanding and generation

Plan-based understanding

PAM

POLITICS

Dynamic Memory

IPP

BORIS

CYRUS

Story telling

TALE-SPIN
Summary

1. Conversational systems have a long history of development.
2. Conversational systems evolved from natural language dialogue systems. Multiple modality and embodiment have been incorporated, then.
3. The Knowledge Navigator gave a significant impact on the concept of intelligent virtual agents.
4. Personality, social interactions and lifelikeness are key features of intelligent virtual agents.
5. In order to create natural conversational behaviors, conversational artifacts need to coordinate verbal and nonverbal behaviors.
6. Story understanding and generation techniques are necessary for processing content.
References


For more details ...