Agent-Oriented Modeling of Real-Life Discourse

GU Yueguo

The Chinese Academy of Social Sciences

guyueguo@vip.sina.com

The Gist

1. Three ways of data-mining: pen-based, audiotaping and videotaping
2. The project: how to extract information from the video streams and construct data-models
3. Long-term goal: to reproduce it through simulation

Main Headings

1. Three ways of data-mining
2. Agent construction
3. Levels of modeling and meta-model diagrams
4. Demonstrations
5. Methodology and tools

1

Three Ways of Data-Mining

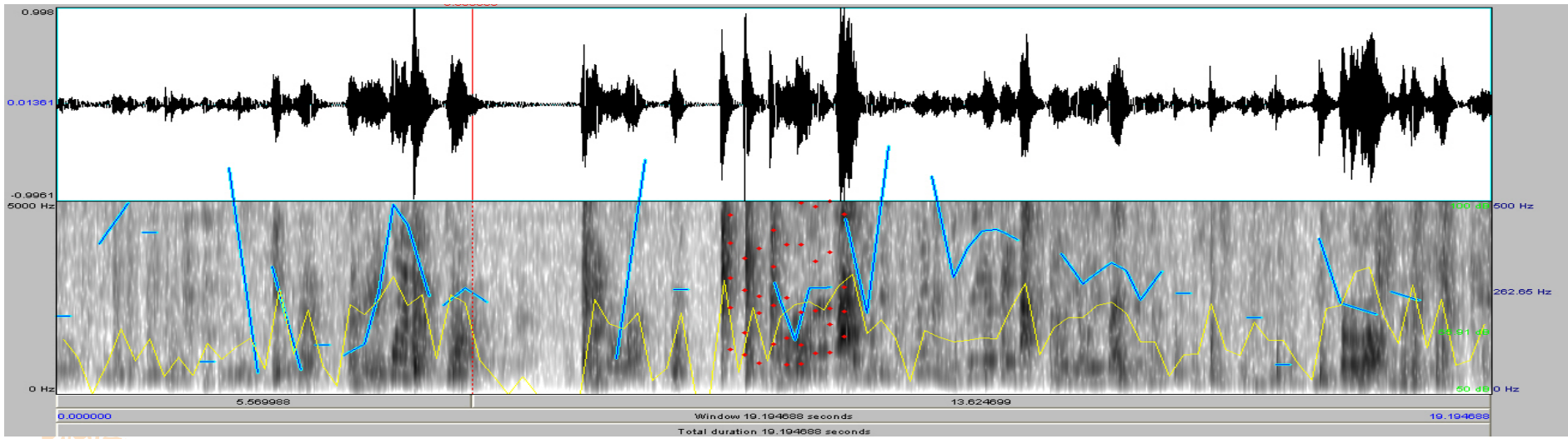
Flux of Real-Life Events

Orthographic representation

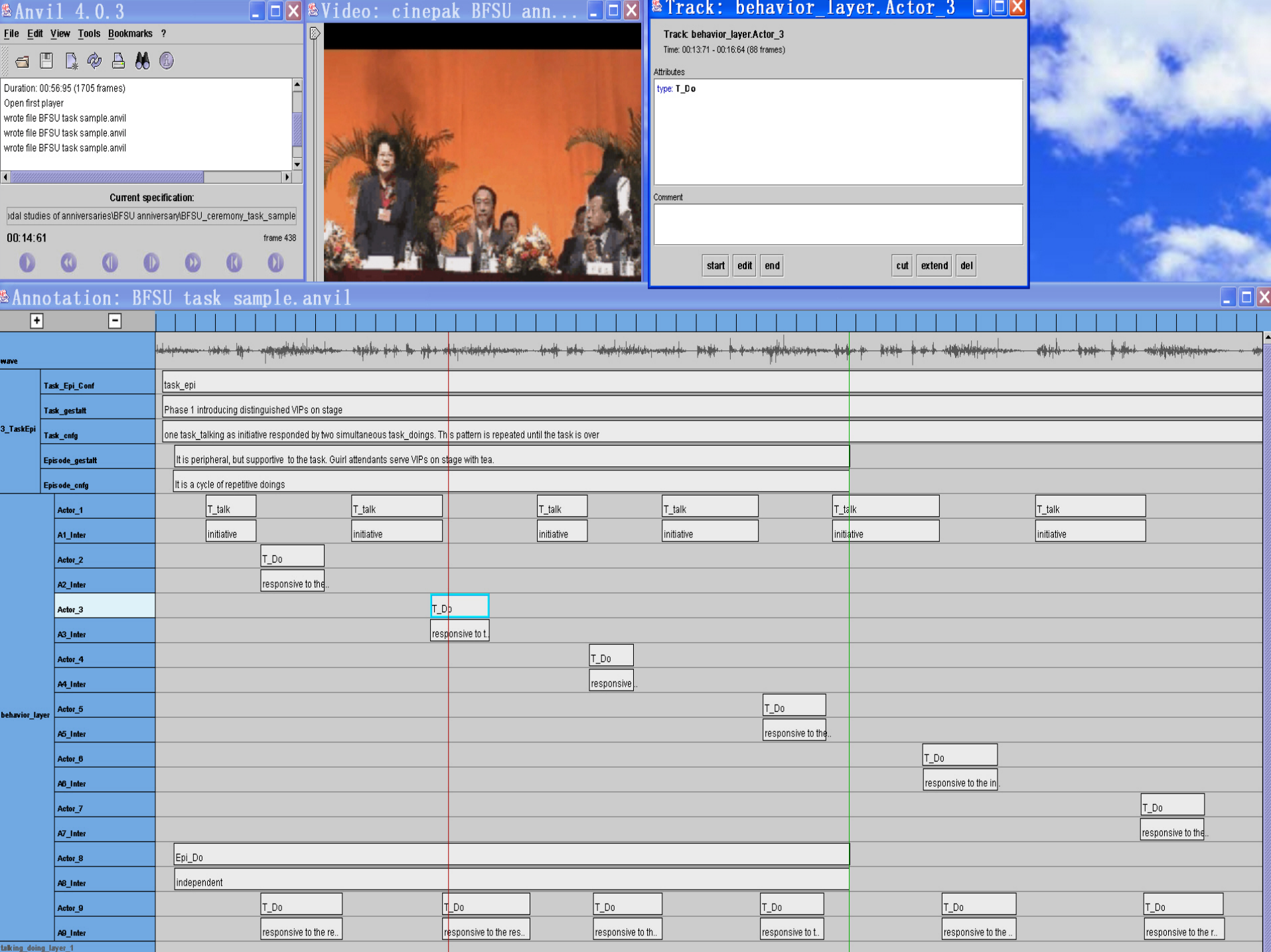
1. 洁：哎?
2. 姥：来呀！……
3. 洁：过来吧！
4. 爷：[咳嗽] 怎么啦？
5. [咳嗽]
6. 大乖乖！
7. 不错！
8. 洁：[拍手]
9. 爷：大乖乖！
10. 慕：看我的大鼻子! 啊！
11. 你声音太大，
12. 所以容易……
13. 爷：大乖乖！
14. 来呀！

Acoustic representation

The audio text of the same verbal content, 19.194688 seconds long, looks like this printed in the paper medium. (The sound file is processed by using Praat.)



Video-streams


Anvil 4.0.3
File Edit View Tools Bookmarks ?
Duration: 00:56:95 (1705 frames)
Open first player
wrote file BFSU task sample.anvil
wrote file BFSU task sample.anvil
wrote file BFSU task sample.anvil
Current specification:
)dal studies of anniversaries\BFSU anniversary\BFSU_ceremony_task_sample
00:14:61
frame 438
Video: cinepak BFSU ann...
Track: behavior_layer.Actor_3
Track: behavior_layer.Actor_3
Time: 00:13:71 - 00:16:64 (88 frames)
Attributes
type: T_Do
Comment
start edit end
cut extend del
Annotation: BFSU task sample.anvil
wave
Task_Epi_Conf
task_epi
Task_gestalt
Phase 1 introducing distinguished VIPs on stage
3_TaskEpi
Task_cnfg
one task_talking as initiative responded by two simultaneous task_doings. This pattern is repeated until the task is over
Episode_gestalt
It is peripheral, but supportive to the task. Guirl attendants serve VIPs on stage with tea.
Episode_cnfg
It is a cycle of repetitive doings
Actor_1
T_talk
A1_Inter
initiative
Actor_2
T_Do
A2_Inter
responsive to the..
Actor_3
T_Do
A3_Inter
responsive to t..
Actor_4
T_Do
A4_Inter
responsive..
behavior_layer
Actor_5
T_Do
A5_Inter
responsive to the..
Actor_6
T_Do
A6_Inter
responsive to the in..
Actor_7
T_Do
A7_Inter
responsive to the..
Actor_8
Epi_Do
A8_Inter
independent
Actor_9
T_Do
A9_Inter
responsive to the re..
responsive to the res..
responsive to th..
responsive to t..
responsive to the ..
responsive to the r..
talking_doing_layer_1

An overview of a piece of real-life situated discourse represented in three different ways via different tools and media

Primary research tools	Representation via medium	Representation texts	Advanced research tools	Digitalized representation text	
	paper/pen-based representation	orthographic text	keyboarding or OCR	digitalized orthographic text	mono-modal, static, orthographic- linear representation in a straight jacket of writing syntax
	audio-based representation	magnetic audio text in a cassette	special applications	digitalized magnetic audio text digitalized acostic text	audio-modal, dynamic/static, time-linear representation in a straight jacket of talking syntax
	video-based representation	magnetic image text in a VHS	special applications	digitalized moving image text digitalized frame -by-frame text	multi-modal, dynamic/static, time-linear representation in a straight jacket of interaction syntax

Three representations contrasted in terms of information loss and trade-offs

Contrasted from the researcher & interpreter's perspective

Mono-modal, static, orthographic- linear representation in a straight jacket of writing syntax -- single	**audio-modal, dynamic/static, time-linear representation in a straight jacket of talking syntax**	**multi-modal, dynamic/static, time-linear representation in a straight jacket of interaction syntax**
Information-losing	**Information-losing compared**	**Information-losing compared:**
Situatedness:	**Situatedness:**	**Situatedness:**
Behavioral setting: complete loss	Behavioral setting: heavy loss	**Behavioral setting: mostly reserved**
The on-going activities: disconnected	The on-going activities: disconnected	The on-going activities: disconnected
Situational goals: sterilized, lost, becoming opague	Situational goals: sterilized, lost, becoming opague	**Situational goals: alive, easily to be reconstructed;**
Situational goal-attaining schema: extremely difficult to reconstruct	Situational goal-attaining schema: less difficult to reconstruct	**Situational goal-attaining schema: fairly easy to be reconstructed;**
De-individualization:	De-individualization:	De-individualization:
(a) sterilize consciousness;	(a) sterilize consciousness;	(a) sterilize consciousness;
(b) cut off life-history;	(b) partially cut off life-history;	**(b) somehow cut off life-history;**
(c) stereotype personality;	(c) partially stereotype personality;	**(c) somehow stereotype personality;**
(d) lose performance style;	(d) partially lose performance style;	**(d) retain performance style;**
(e) sterilize and stereotype roles	(e) partially sterilize and stereotype roles	**(e) somehow sterilize and stereotype roles**
The intersubjective world -- extremely difficult to reconstruct;	The intersubjective world -- less difficult to reconstruct;	**The intersubjective world -- much less difficult to reconstruct;**
Interdependency relations: natural ties mostly lost;	Interdependency relations: natural ties being loosened or even lost;	**Interdependency relations: natural ties easily reconstructable;**

Continued

audio information: complete loss	audio information: retained	audio information: retained
individualized voice quality	individualized voice quality	individualized voice quality
affective features	affective features	affective features
prosodic features	prosodic features	prosodic features
segmental features	segmental features	segmental features
Video information: complete loss	**Video information: complete loss**	**Video information: retained**
behavioral	behavioral	behavioral
kinetic	kinetic	kinetic
spatial	spatial	spatial
Authenticity: complete loss	**Authenticity: partial loss**	**Authenticity: retained**
Validation: extremely difficult	**Validation: less difficult**	**Validation: easy**
Trade-off benefits	**Trade-off benefits compared:**	**Trade-off benefits compared:**
Highly discretized	voice streams, discretization now much easier, but still costly	image streams,discretization still difficult and costly
Lexicalized (conceptualized) perception (vs. analog perception)	a lot of features difficult to be lexicalized, or even non-lexicalizable;	a lot of features difficult to be lexicalized, or even non-lexicalizable;
Highly friendly to abstract thinking	less friendly to abstract thinking, but friendly to musical thinking;	less friendly to abstract thinking, but friendly to gestalt thinking;
Easily processible	still difficult to process	very difficult to process

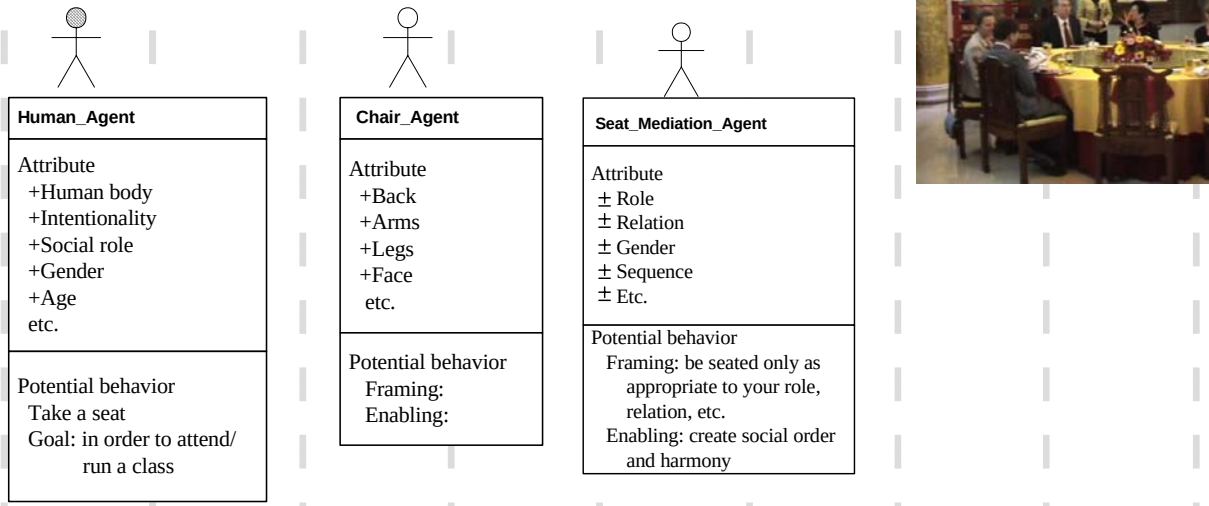
2

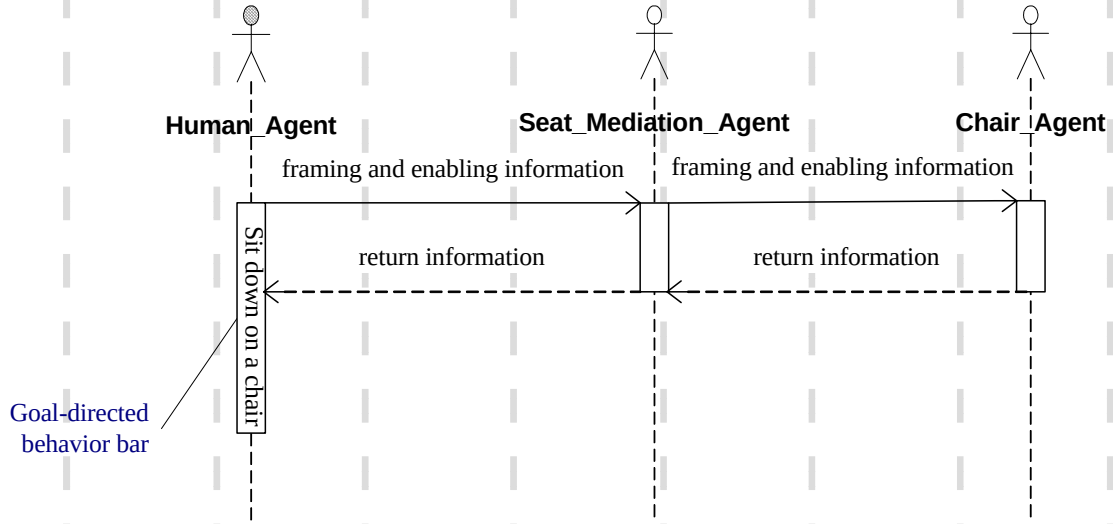
Agent Construction

Event-reporting vs. process-modeling

Event reporting: John took the leftist seat.

Process narrative: John appeared in the banquet room. He looked around and studied the layout of the table. He reflected upon his own position and decided that the leftist seat is appropriate to his status.





A brief review

As we know, the notion of agent construction recently has been intensely explored in artificial intelligence (see e.g. Hexmoor, Castelfranchi, and Falcone 2003, Alonso, Kudenko and Kazakov, 2003, Ye and Churchill 2003, Russell and Norvig 2003, Wagner 2004).

Russell and Norvig 2003

An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators. (p.32; bold original)

... an agent is meant to be a tool for analyzing systems, not an absolute characterization that divides the world into agents and non-agents. (p.34)

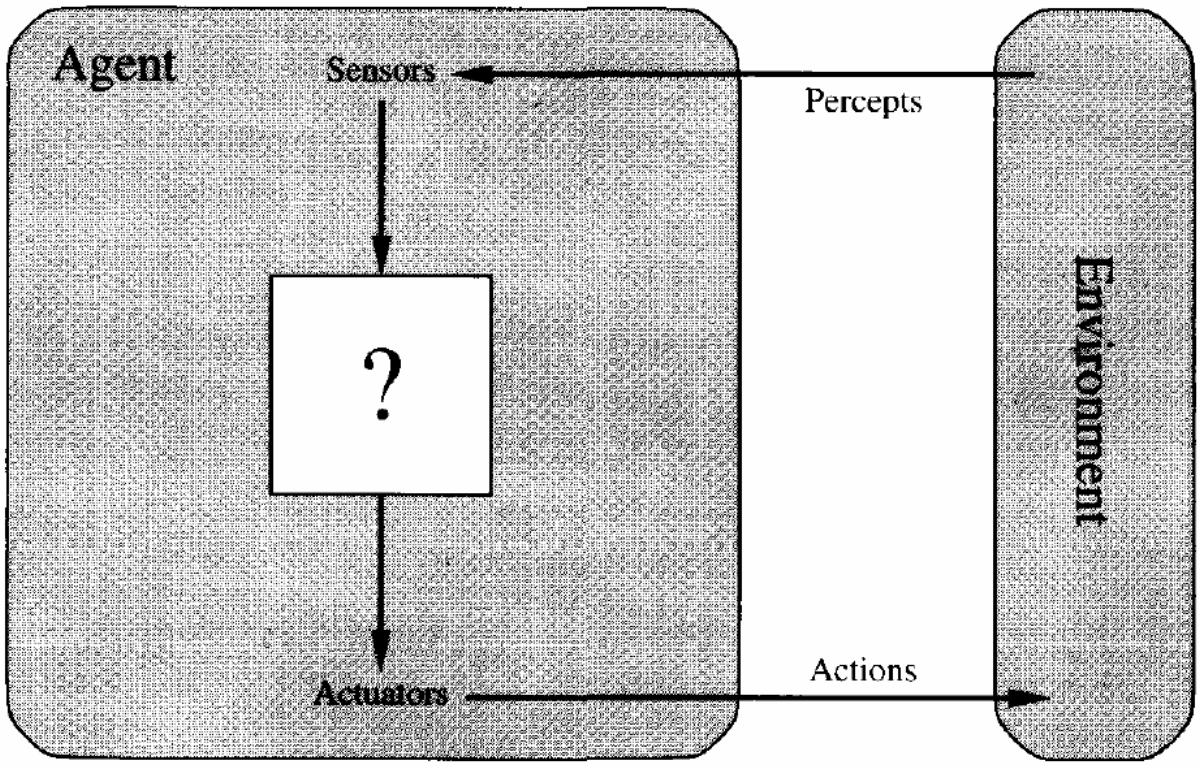


Figure 2.1 Agents interact with environments through sensors and actuators.

… we have talked about agents by describing *behavior* --- the action that is performed after any given sequence of percepts. … The job of AI is to design the **agent program** that implements the agent function mapping percepts to actions. We assume this program will run on some sort of computing device with physical sensors and actuators --- we call this the **architecture**:

agent = architecture + program

(Russell and Norvig 2003: 44; italics and bold original)

Two Basic Descriptors

An agent will have two basic descriptors: (1) attribute and (2) behavior.

Attributes are properties abstracted from the object of investigation *that are pertinent to the interaction of the agents involved*. For instance, we are not interested in all the attributes that a chair potentially possesses. We only abstract those attributes that are pertinent to the approaching human agent (i.e. a human guest) who is pursuing the goal of sitting down on it

attributes of agents are abstracted on the basis of the activity type and the goals that are being pursued by the agents (human or non-human).

AML agents vs. AI agents

- As a tool for abstracting data from the real world
- The real world is seen as a network of interactions between all sorts of agents
- Agents' behavior/action is an abstraction from the real-world event/process, having little to do with programming

- As a tool for analyzing systems
- There is no such ontological relation
- Agent's behavior/action is a program operating on a particular architecture

Types of AI Agents (Russell and Norvig 2003)

1. Simple reflex agents
2. Model-based reflex agent
3. Model-based, goal-based agent
4. Model-based, utility-based agent

Differences in the conceptualization of agents

- Although I have drawn a great deal of inspiration from the AI literature, the *agent* concept dealt with in this paper should not be regarded as an application of that defined by AI researchers.
- I look at the AI agent as part of a programming metalanguage, while the agent in this paper is intended to be part of a modeling metalanguage.
- I owe more debts to the literature on the OOP and the UML (unified modeling language) than to the AI agent (see the next subsection for detailed discussion).

The Agent-based Modeling Language (AML)

The primary purpose of using the AML to model the real world activity is to build a structured representation of data which is otherwise elusive and resists information engineering. So the AML must meet two basic requirements:

(1) mine the data from real life activities, and

(2) construct the extracted data in such a way that the data can easily be manipulated by computing.

3

Levels of Modeling and Meta Diagrams

Four Levels of Modeling

1. *Instance Modeling.*
2. *Activity Type Modeling.*
3. *Applied Modeling.*
4. *Schema and Database Modeling.*

Packaging and meta-model diagrams

1. behavioral and dynamic package;
2. structural and static package, and
3. application design package.

4

A Demo

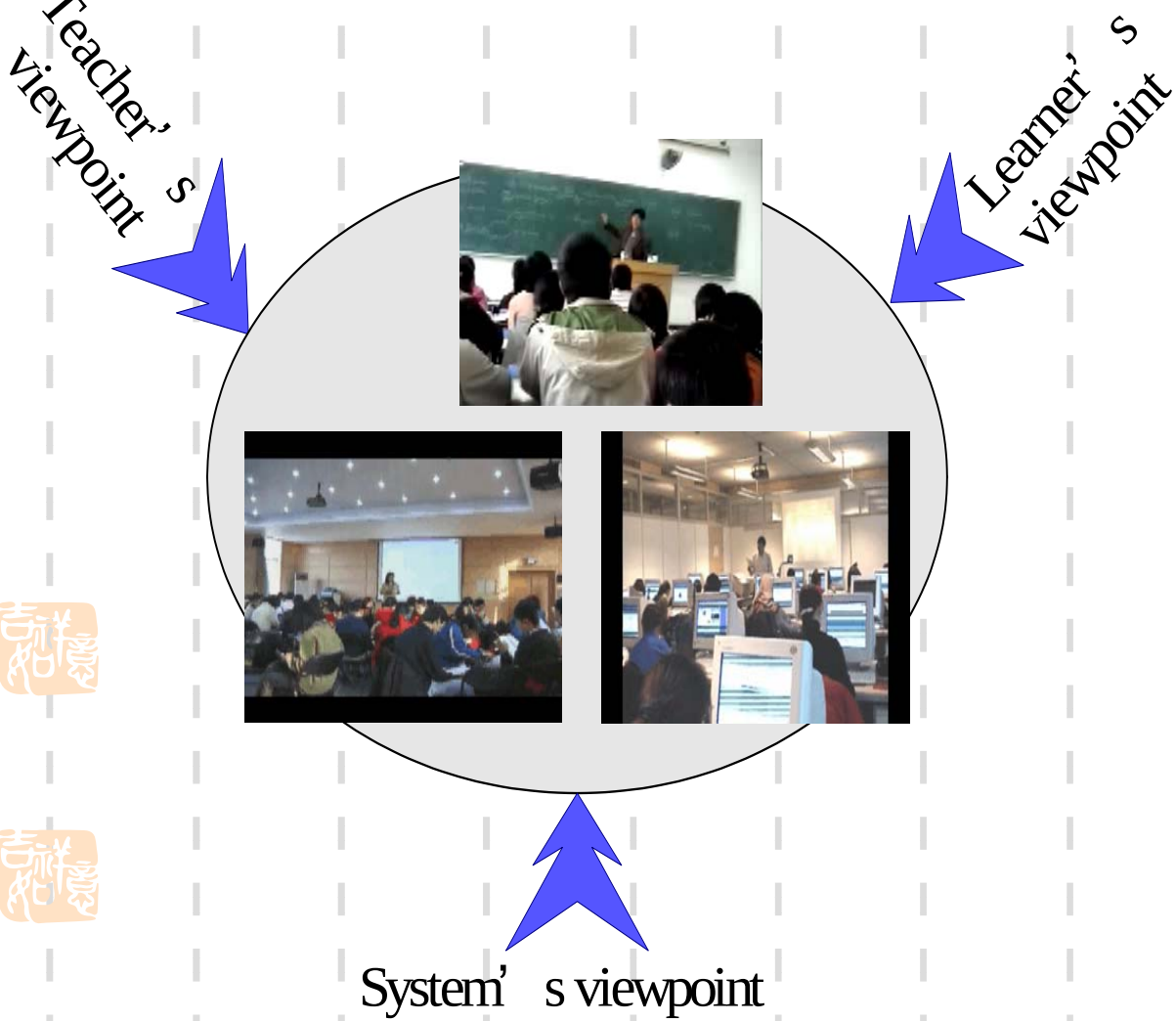


Figure xx: Three complementary views of a classroom discourse

A demo

1. ***The teacher's instructional talk/do diagram***: a sequence of a teacher's behavior (e.g. an hour of teaching) segmented and annotated in terms of talking and doing. As pointed out in xxx above in this paper we deal with behavior mainly at molar level, that is, we shall not further categorize talking/doing into sub-categories (but see an exception of this general restriction at xxx below). The same will be true of the learners' talking and doing;
2. ***The learners' learning talk/do diagram***: a sequence of learners' behavior (e.g. an hour of learning with a teacher or with a CD-ROM, or a Web-based courseware) segmented and annotated in terms of talking and doing;
3. ***The teacher's interaction diagram***: patterns of the teacher's interactions with learners, the PPT, the blackboard, the computer screen, and so on;
4. ***The learners' interaction diagram***: patterns of the learners' interactions with the teacher, fellow learners, cybermates, print-texts, and so on;
5. ***The teacher's goal-attaining process diagram***: the teacher's sequence of instructional behavior segmented and annotated in terms of instructional tasks performed to achieve instructional goals/sub-goals;
6. ***The learners' goal-attaining process diagram***: the learner's sequence of learning behavior segmented and annotated in terms of learning tasks performed to achieve learning goals;
7. ***The session's gestalt diagram***: an integrated diagram showing both the teaching and the learning behaviors.

1 The teacher's talk/do diagram

Figure xx: the teacher's talk/do diagram

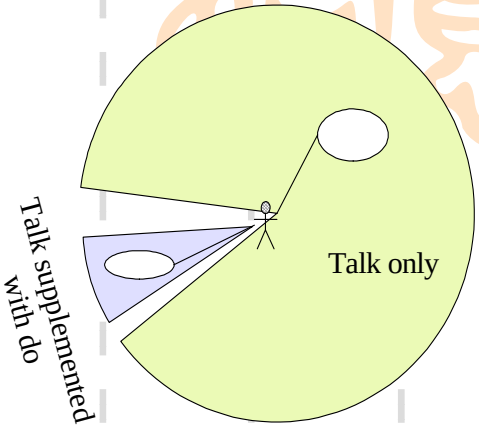


Figure xx: Videotaped classroom teaching

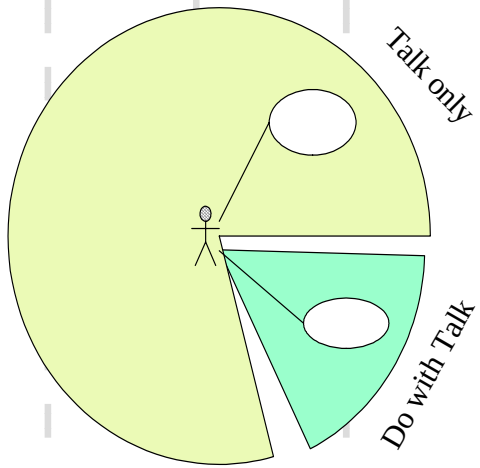


Figure xx: the teacher's talk/do diagram in a grammar class

2. The learner's talk/do diagram

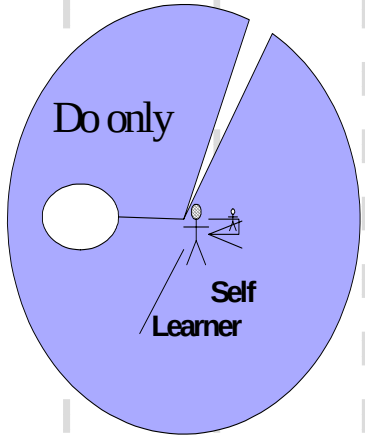


Figure xx: the learner's talk/do diagram in the grammar class

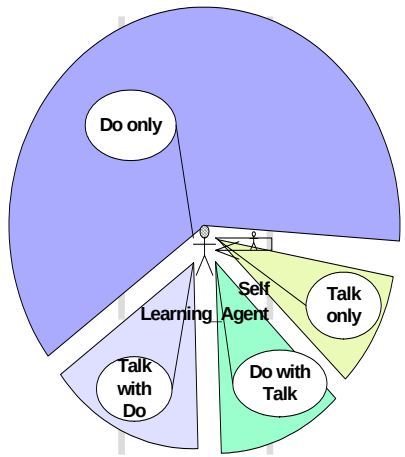


Figure xx: the learner's talk/do diagram in the English skill class

3. The teacher's interaction diagram

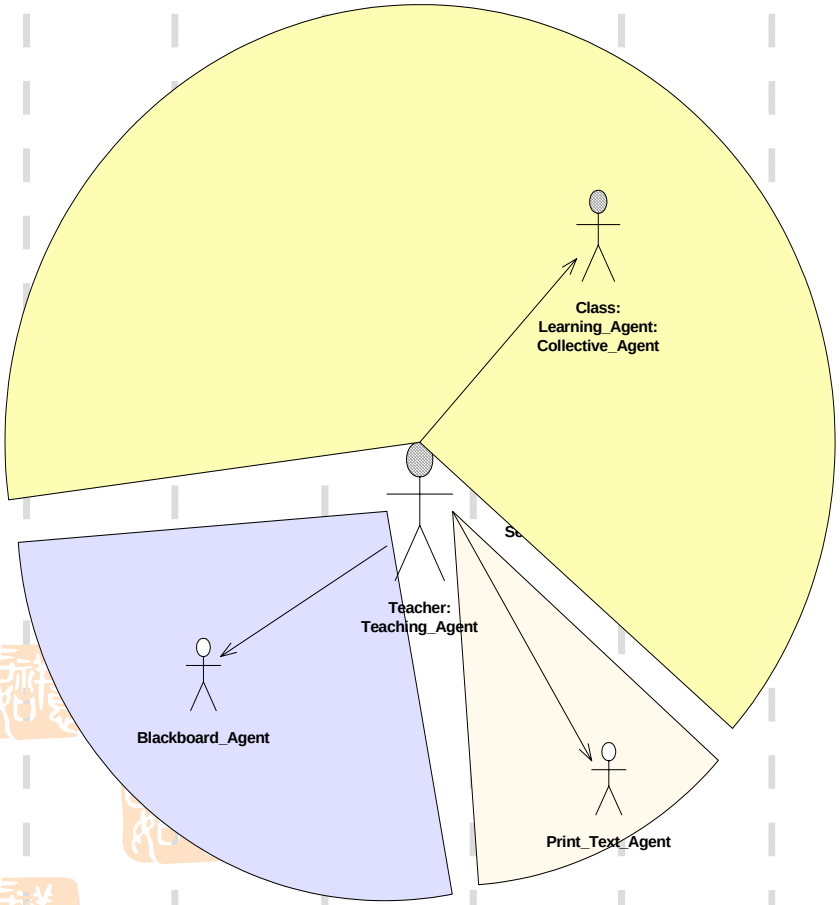


Figure xx-a: the grammar teacher's interaction diagram

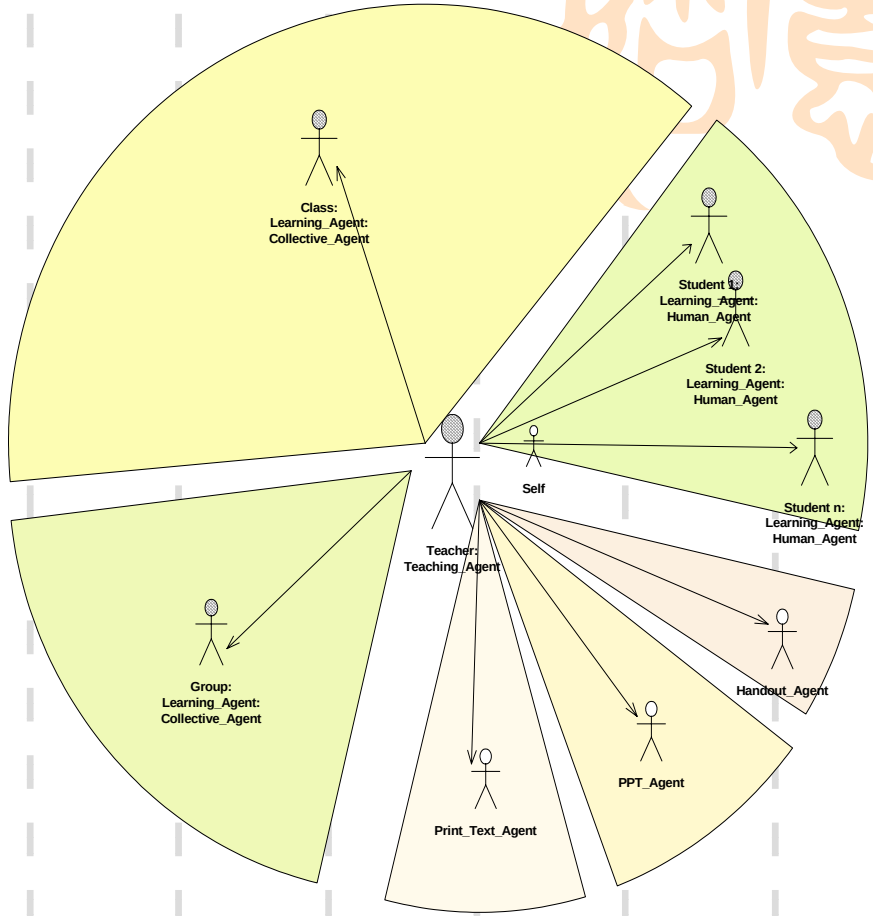


Figure xx-b: the literature teacher's interaction diagram

4. The learner's interaction diagram

Figure xx: The learner's interaction diagram in the literature class

5. The teacher's goal attaining process

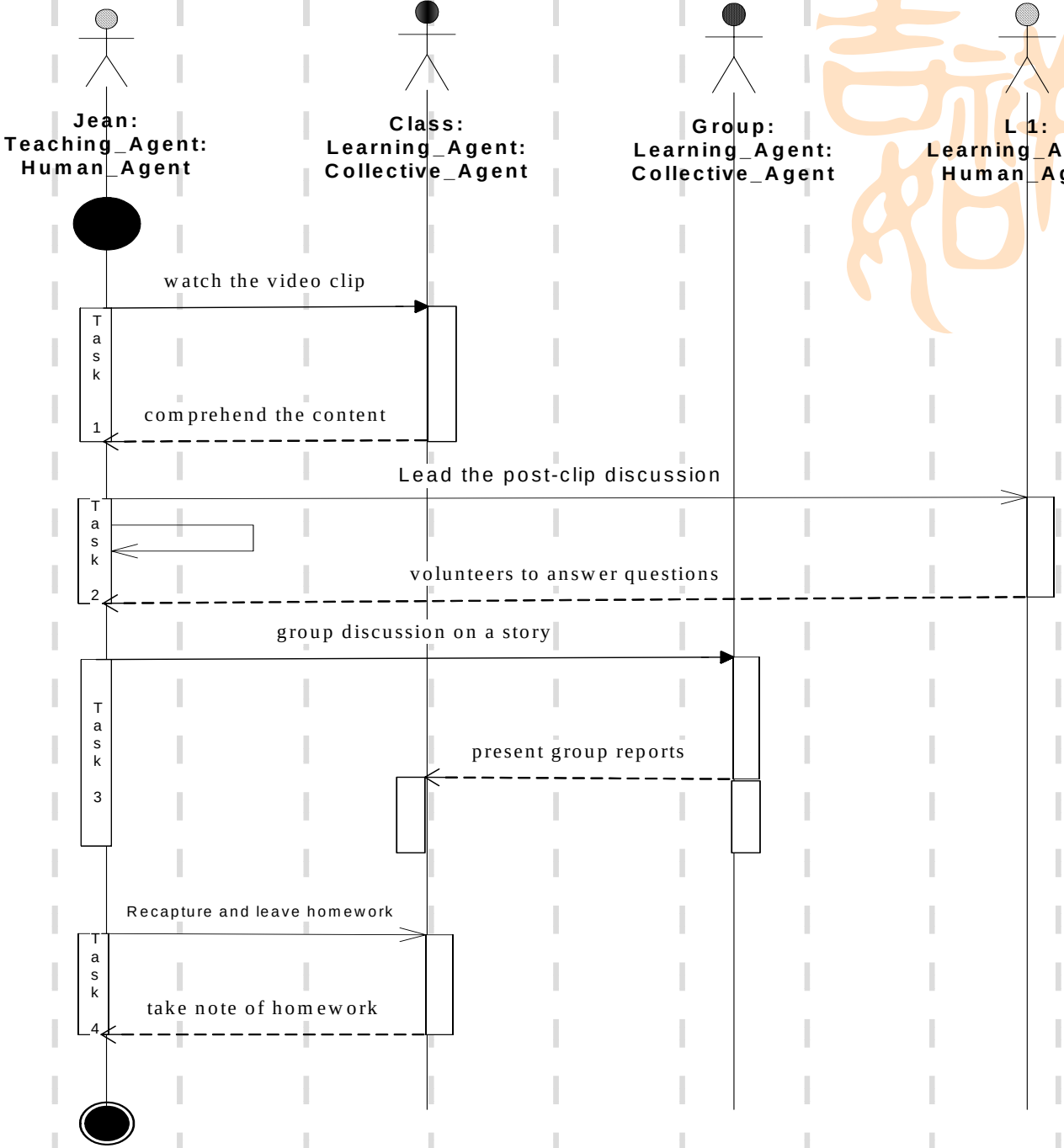
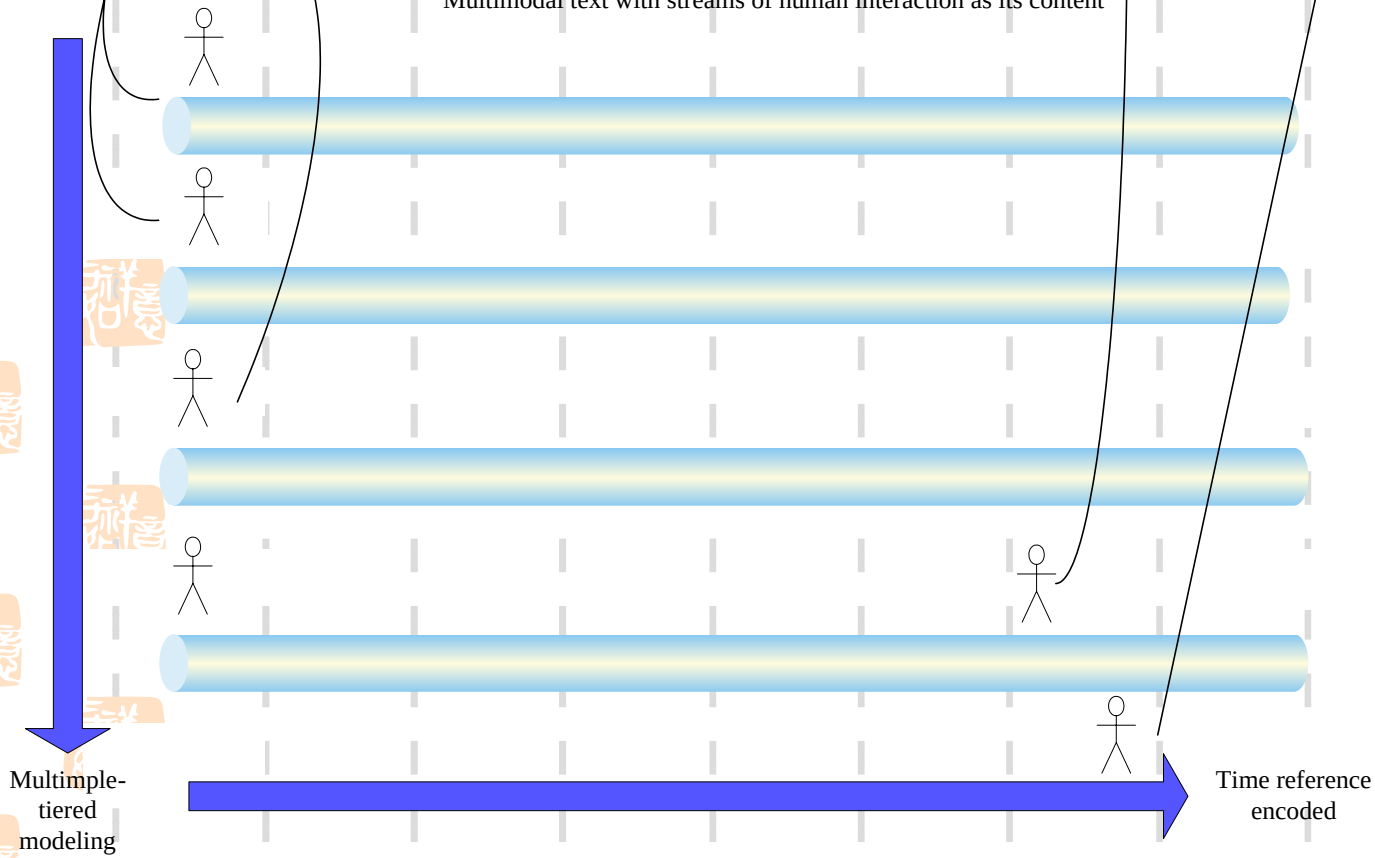
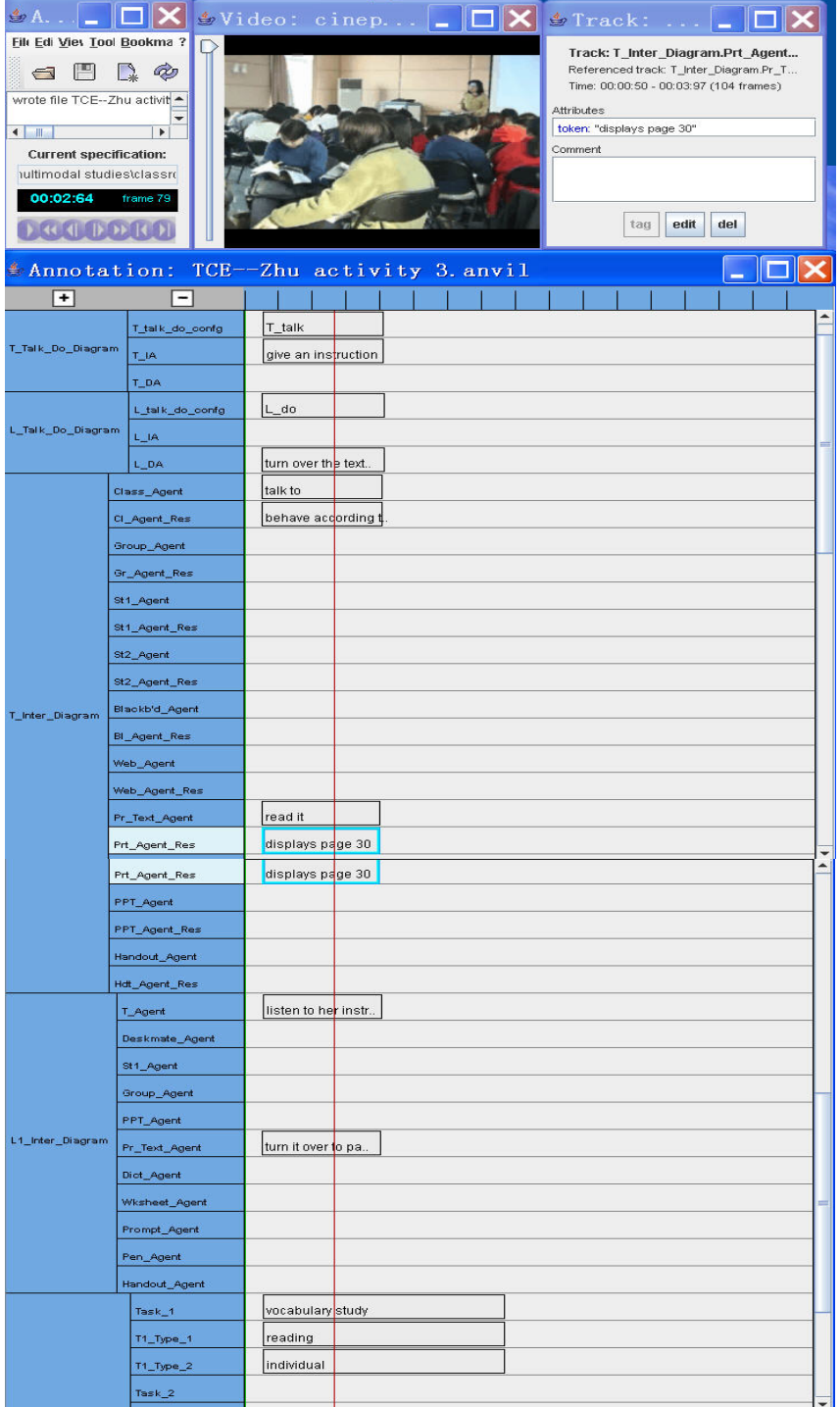


Figure xx: The literature teacher's goal-attaining process diagram

7. The session's gestalt diagram

Multimodal text with streams of human interaction as its content




Video: cinep...
Track: ...
Track: T_Inter_Diagram.Prt_Agent...
Referenced track: T_Inter_Diagram.Pr_T...
Time: 00:00:50 - 00:03:97 (104 frames)
Attributes
token: "displays page 30"
Comment
tag
edit
del
Current specification:
00:02:64
frame 79
Annotation: TCE--Zhu activity 3.anvil
T_Talk_Do_Diagram
T_talk_do_config
T_IA
T_DA
T_talk
give an instruction
L_Talk_Do_Diagram
L_talk_do_config
L_IA
L_DA
L_do
turn over the text..
T_Inter_Diagram
Class_Agent
talk to
Cl_Agent_Res
behave according t..
Group_Agent
Gr_Agent_Res
St1_Agent
St1_Agent_Res
St2_Agent
St2_Agent_Res
Blackb'd_Agent
Bl_Agent_Res
Web_Agent
Web_Agent_Res
Pr_Text_Agent
read it
Prt_Agent_Res
displays page 30
Prt_Agent_Res
displays page 30
PPT_Agent
PPT_Agent_Res
Handout_Agent
Hdt_Agent_Res
L1_Inter_Diagram
T_Agent
listen to her instr..
Deskmate_Agent
St1_Agent
Group_Agent
PPT_Agent
Pr_Text_Agent
turn it over to pa..
Dict_Agent
Wksheet_Agent
Prompt_Agent
Pen_Agent
Handout_Agent
Task_1
vocabulary study
T1_Type_1
reading
T1_Type_2
individual
Task_2

System's behavior diagrams

human political geography

- control of ideology
- regional power distribution
- control of mass media
- control of policy-making

Political, and ideological space and order are carried out

human social geography

- social stratification
- social networks
- social status
- social discrimination
- access to education
- access to healthcare

Social geography frames and at the same time enables the formation of political geography

human economic geography

5 Quinary activities
- Executive decision making

4 Quaternary activities
- information
- research
- management

3 Tertiary activities
- retail and wholesale trade
- personal and professional services

2 Secondary activities
- manufacturing
- processing
- construction
- Power production

1 Primary activities
- agriculture
- gathering industries
- extractive industries

Social initiatives, social maintenance, social order, etc. are carried out

Economic geography frames and at the same time enables the formation of social geography

habitat geography

Economical activities are carried out in local habitas and globally

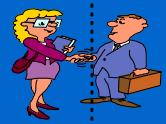
Habitat geography frames and at the same time enables human activities

Physical geography

Humans adapt themselves to it, and transform it, and migrating over it

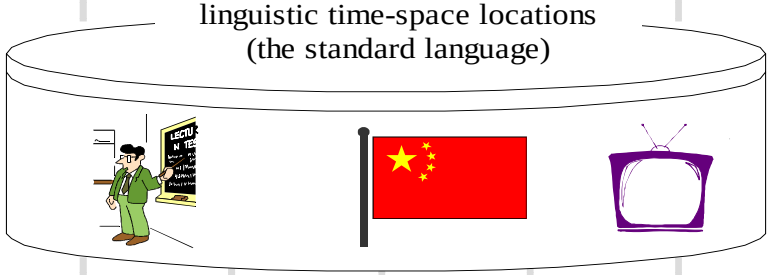
The physical environment frames and at the same time enables human settlements

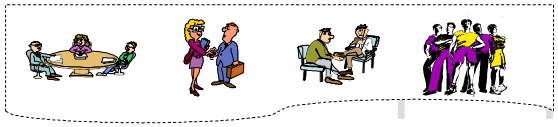
Observable, describable collective activities



Private and personal behavior

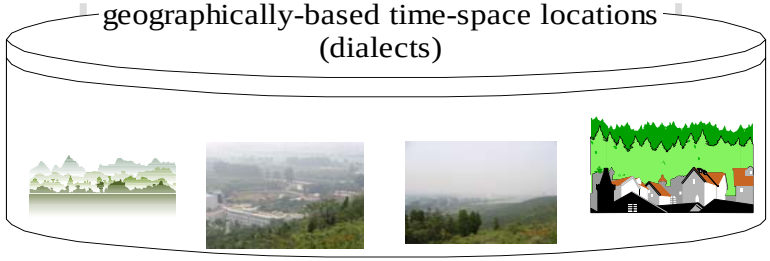
a web of life-paths of a population over the specially constructed linguistic time-space locations (the standard language)

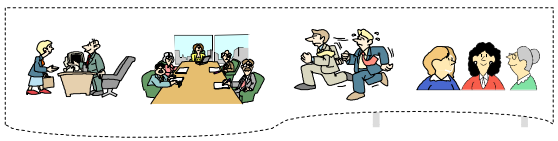




Situated discourses dynamically produce, maintain and reproduce the standard language. They are also an interface of interaction between the standard language, dialects and idiolects.

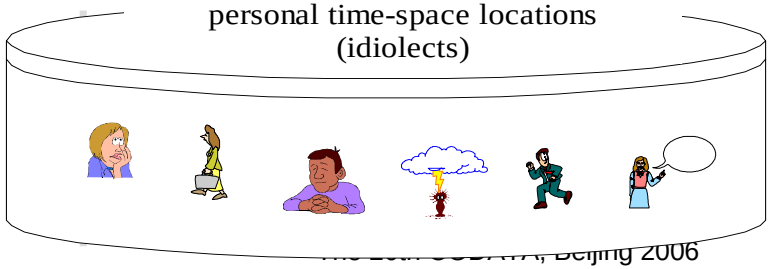
a web of life-paths of a population over geographically-based time-space locations (dialects)

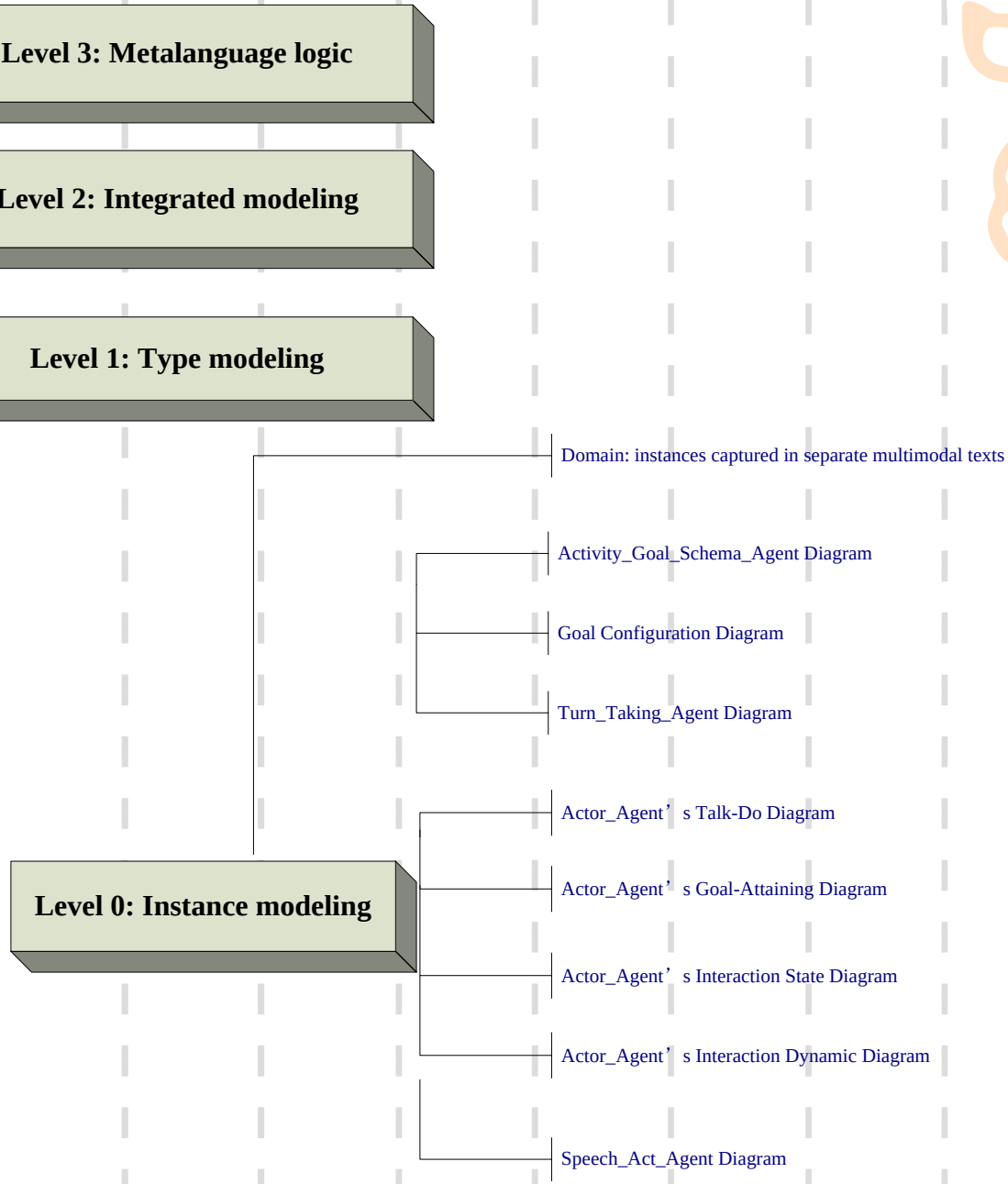




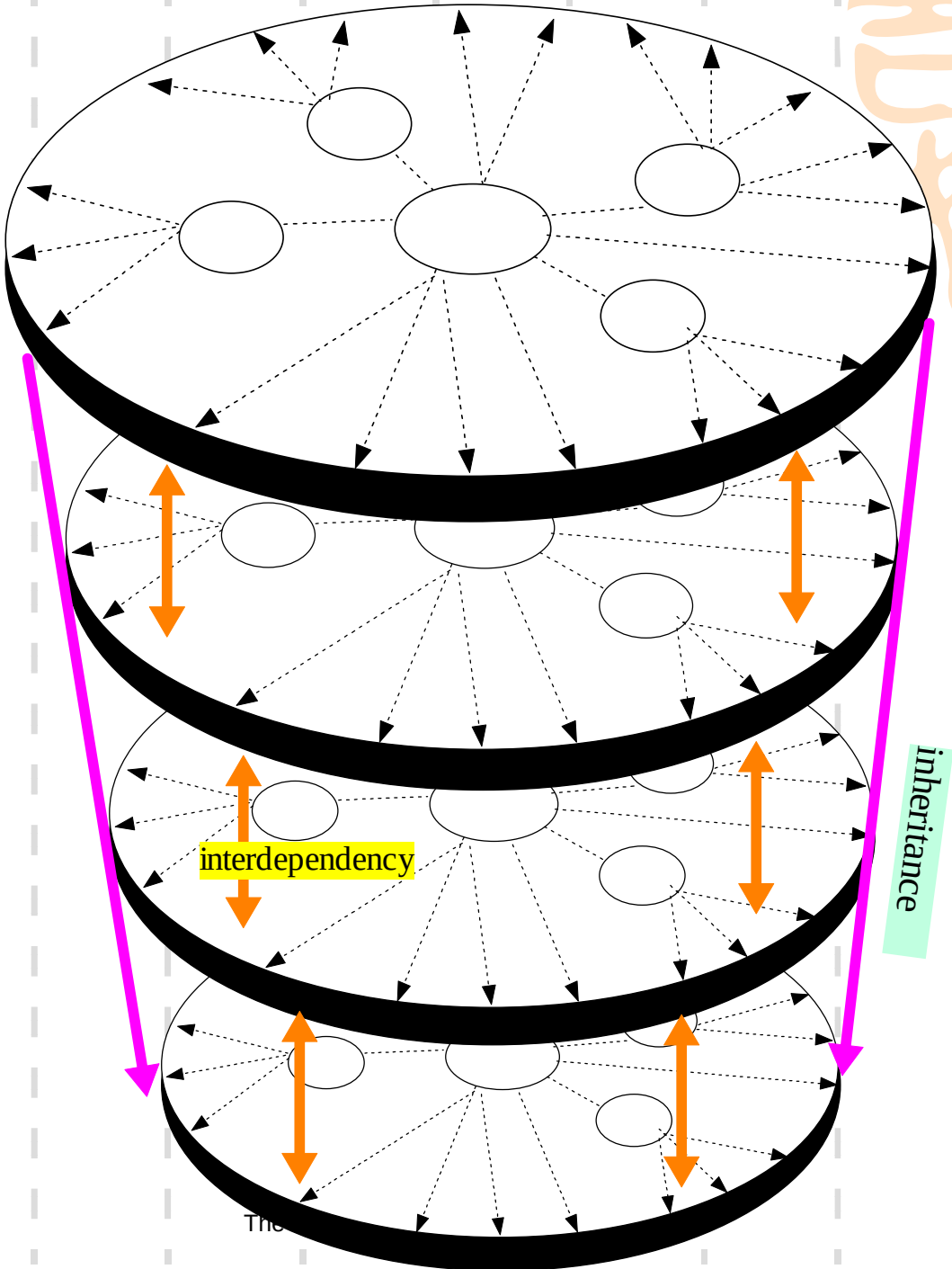
Situated discourses dynamically produce, maintain and reproduce dialects. They are also an interface of interaction between the standard language, dialects and idiolects.

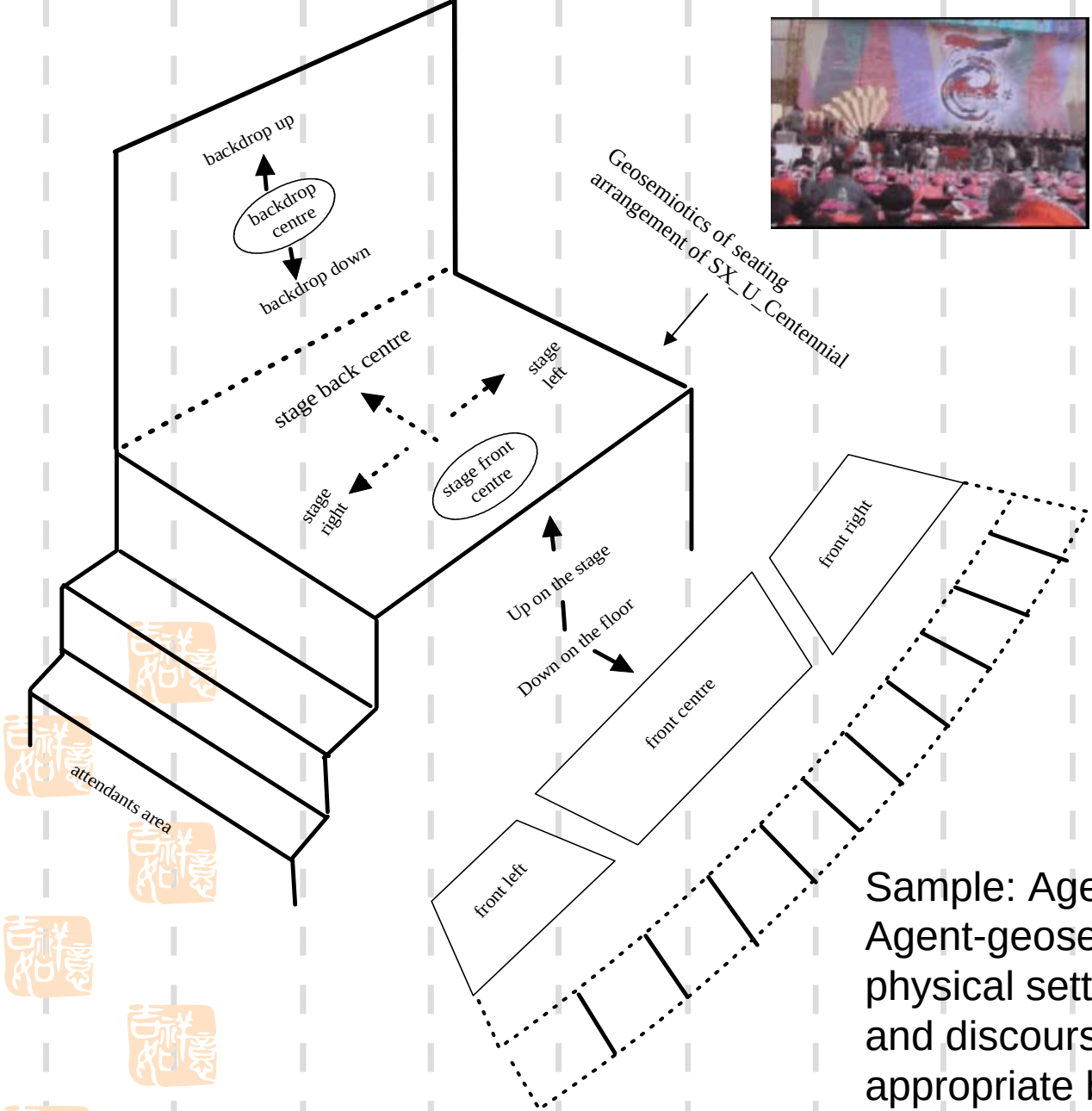
An individual's life-path trajectories over personal time-space locations (idiolects)




Level 3: Metalanguage logic
Level 2: Integrated modeling
Level 1: Type modeling
Level 0: Instance modeling
Domain: instances captured in separate multimodal texts
Activity_Goal_Schema_Agent Diagram
Goal Configuration Diagram
Turn_Taking_Agent Diagram
Actor_Agent’s Talk-Do Diagram
Actor_Agent’s Goal-Attaining Diagram
Actor_Agent’s Interaction State Diagram
Actor_Agent’s Interaction Dynamic Diagram
Speech_Act_Agent Diagram

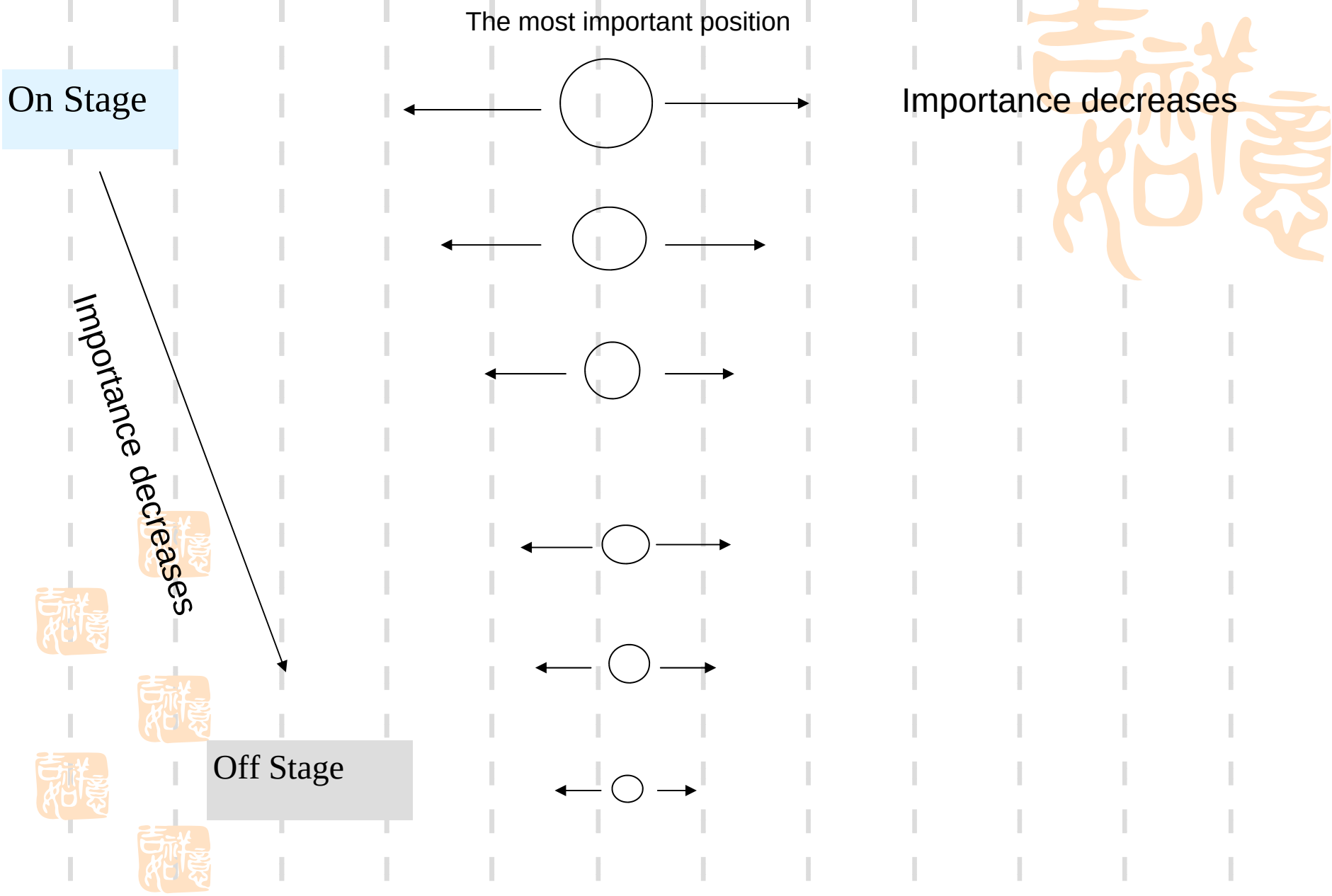
Multi-eco-agent modeling



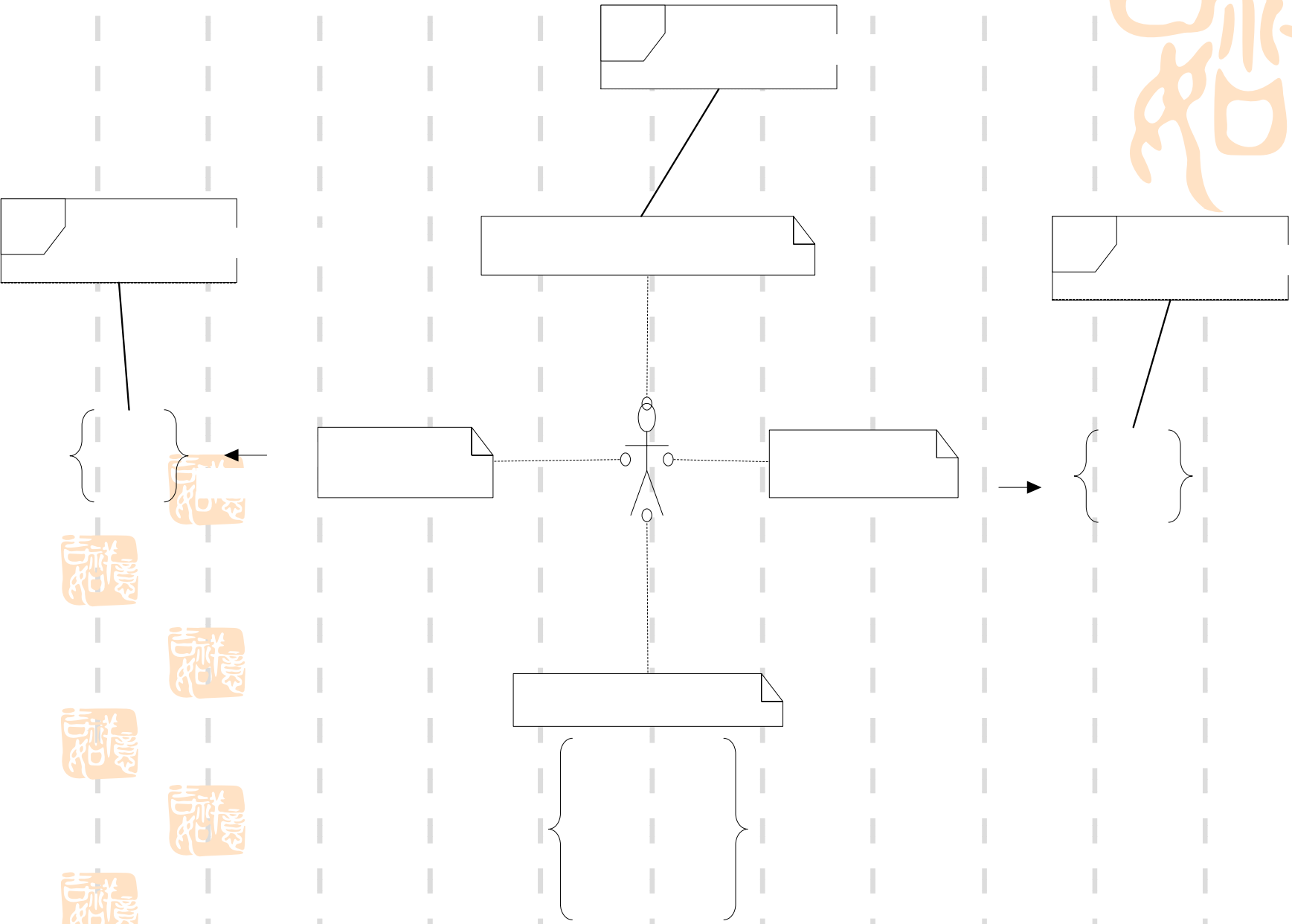


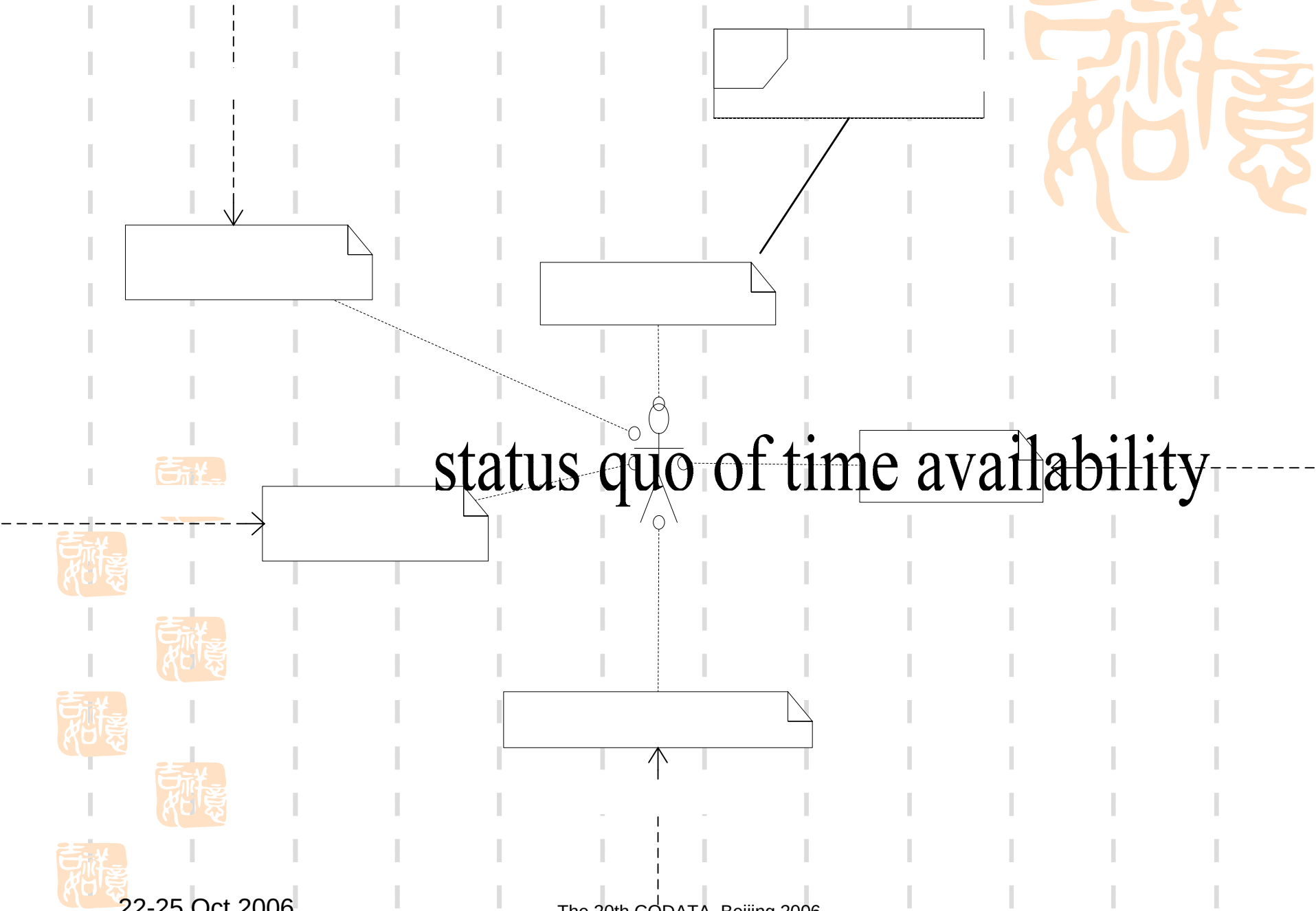
Sample: Agent-beh_setting and Agent-geosemiotic managing physical setting mapped with social and discoursal roles, and generating appropriate behavior patterns

To be continued



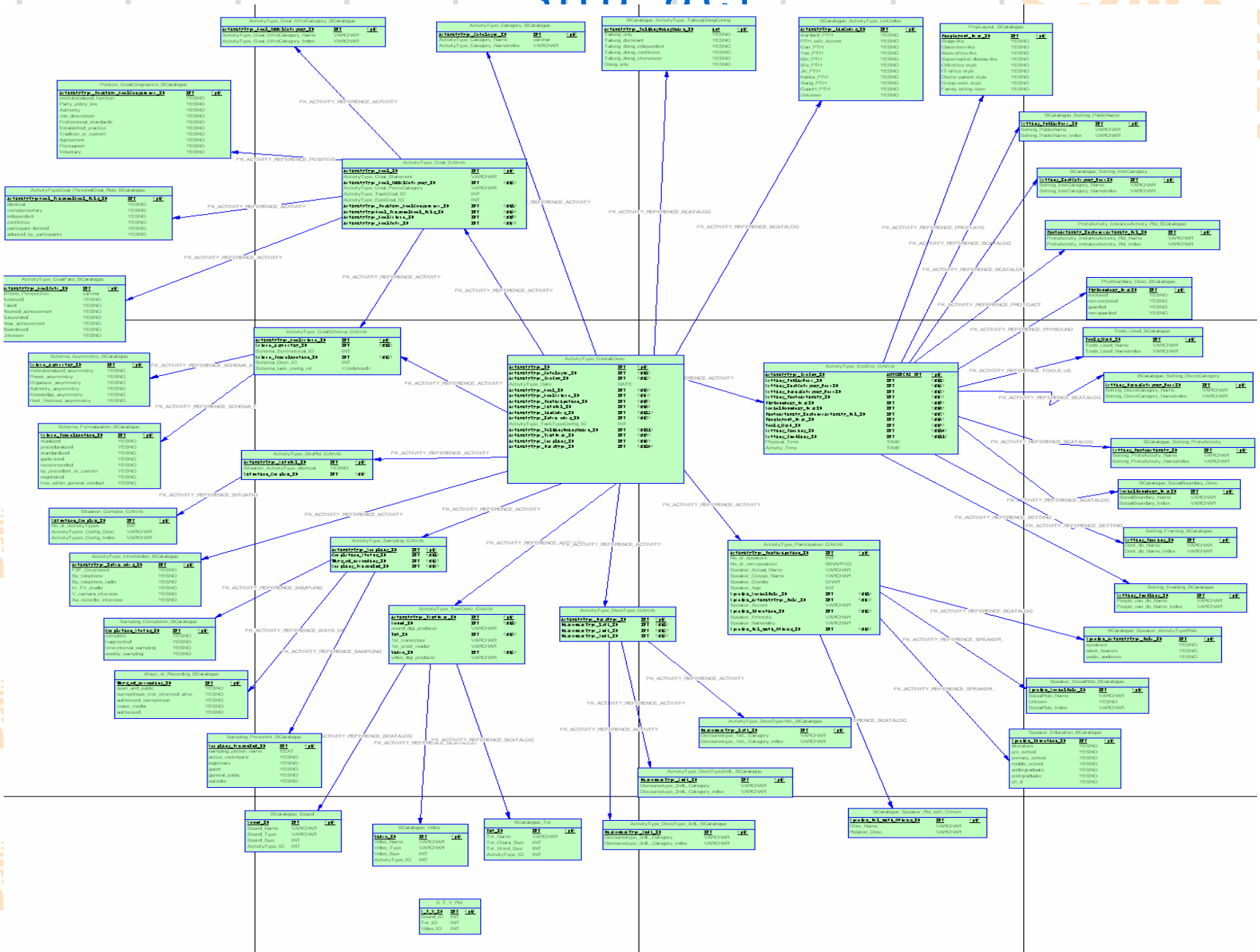
Seating arrangements reflecting the occupants' social importance in Chinese culture. In the present study, the talking and doing behaviors of the occupants inherit the geosemiotic values





Consulting

Database Representation of the Agent-situ act



5

Methodology and tools

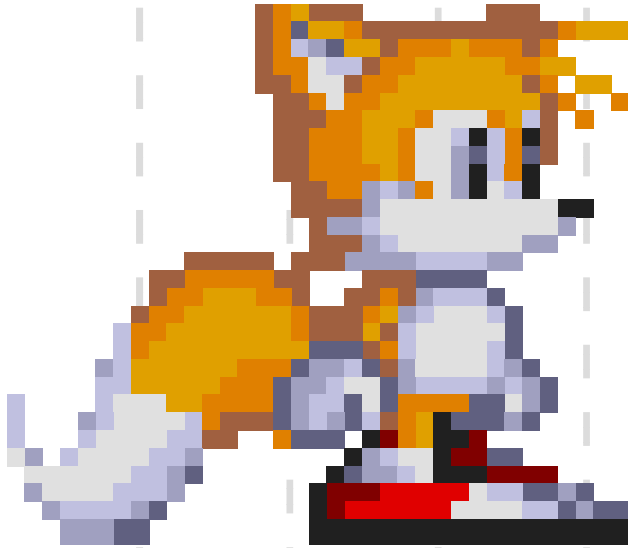
Modeling as a way of solving complexity

建模作为化解复杂为简单的重要方法论

1. Objected Oriented Programming
2. Java
3. UML
4. Agent-oriented design in AI

Tools

1. Agent Card Modeling (ACM)
2. Anvil 4.5
3. UML



Many, many thanks for your participation and contribution!

mentally semiotic nature of language. For example, sociologist Norbert Elias maintained that to understand virtually anything, humans must be able to distance themselves from physical reality: "they must, as it were, mentally ascend to a level of synthesis above that of its existence here and now as a heap of matter."[17] Various types of symbolic representations allow

World-building, Representation and Modeling

1. The ecological relativism: the object of investigation = the animal + perceived environment
   a) The perceiving tool
   b) The perspective
   c) The world-constructing language (cf. Steward's world-building-and-rebuilding 1996:3)
   d) The level of abstraction
   e) The product (oral-aural streams of sounds)
2. The tool-assisted world-understanding (data-mining)
   a) The object of investigation = the animal + tool-assisted perception of the environment
   b) Modeling
      a) The perceiving tool
      b) The perspective
      c) The representation language
   c) The application
3. The digital simulated world (intelligent use of information)
   a) Modeling
      a) The perceiving tool
      b) The perspective
      c) The modeling language
   b) The programming language
   c) The virtual world display language

Ecological relativisim

1. Squirrel's world

1. Human world

1. World through the video camera lens
2. Now-time
3. Here-space
4. Integrated simultaneity
5. parallel

Primary and primitive 1: Naked Personal Experience

1. The object of investigation — flux of events
2. The perceiving tool --- five senses of humans
3. The perspective --- body high, 180 degrees
4. The representation language --- natural speech and “beyond-word mental representation”
5. The level of abstraction --- determined by the natural language used, as well as by the mental representation
6. The product --- mixed memories that can be recalled by the experiencer
7. The application --- for personal experience

Primary and aided 2: hand-drawn graphic representation

1. The object of investigation — flux of events
2. The perceiving tool --- five senses of humans
3. The perspective --- body high, 180 degrees
4. The representation language --- drawings
5. The level of abstraction --- as accurate as possible
6. The product --- hand-drawn pictures that can be looked at by humans later
7. The application --- for sharing with other human beings

Primary and aided 3: Written text representation

1. The object of investigation — flux of events
2. The perceiving tool --- five senses of humans
3. The perspective --- body high, 180 degrees
4. The representation language --- written natural language
5. The level of abstraction --- determined by the writing system
6. The product --- a written text that be read by humans later
7. The application --- for sharing with those who can read the language

Secondary and aided 1: Audio/analogue representation

1. The object of investigation — flux of events
2. The perceiving tool --- an audio analogue recording device
3. The perspective --- distance, and all around
4. The representation language --- magnetic tracks
5. The level of abstraction --- determined by the capability of the device
6. The product --- an audio text that can be listened to later
7. The application --- for sharing audio experience

Secondary and aided 2: Video/analogue representation

1. The object of investigation — flux of events
2. The perceiving tool --- video camera
3. The perspective --- varied angles
4. The representation language --- magnetic tracks
5. The level of abstraction --- determined by the capability of the device
6. The product --- a video text that can be watched later by human beings
7. The application --- for personal experience

Digital era 1: Written text processing

1. The object of investigation — flux of events
2. The perceiving tool --- inherited
3. The perspective --- inherited
4. The representation language --- machine language to re-represent the natural language written text
5. The level of abstraction --- determined by the machine language, as well as by the natural language
6. The product --- electronic text that can be manipulated by the computer
7. The application --- for information processing and sharing

Digital era 2: audio text processing

1. The object of investigation — flux of events
2. The perceiving tool --- inherited
3. The perspective --- inherited
4. The representation language --- machine language to re-represent the analog information
5. The level of abstraction --- determined by the machine language, as well as by the analog representation
6. The product --- electronic audio text that can be manipulated by the computer
7. The application --- for information processing and sharing

Digital era 3: Video text editing

1. The object of investigation — flux of events
2. The perceiving tool --- inherited
3. The perspective --- inherited
4. The representation language --- machine language to re-represent the analog information
5. The level of abstraction --- determined by the machine language, as well as by the analog representation
6. The product --- electronic video text that can be manipulated by the computer
7. The application --- for information processing and sharing

Neural representation?

Psychological representation?

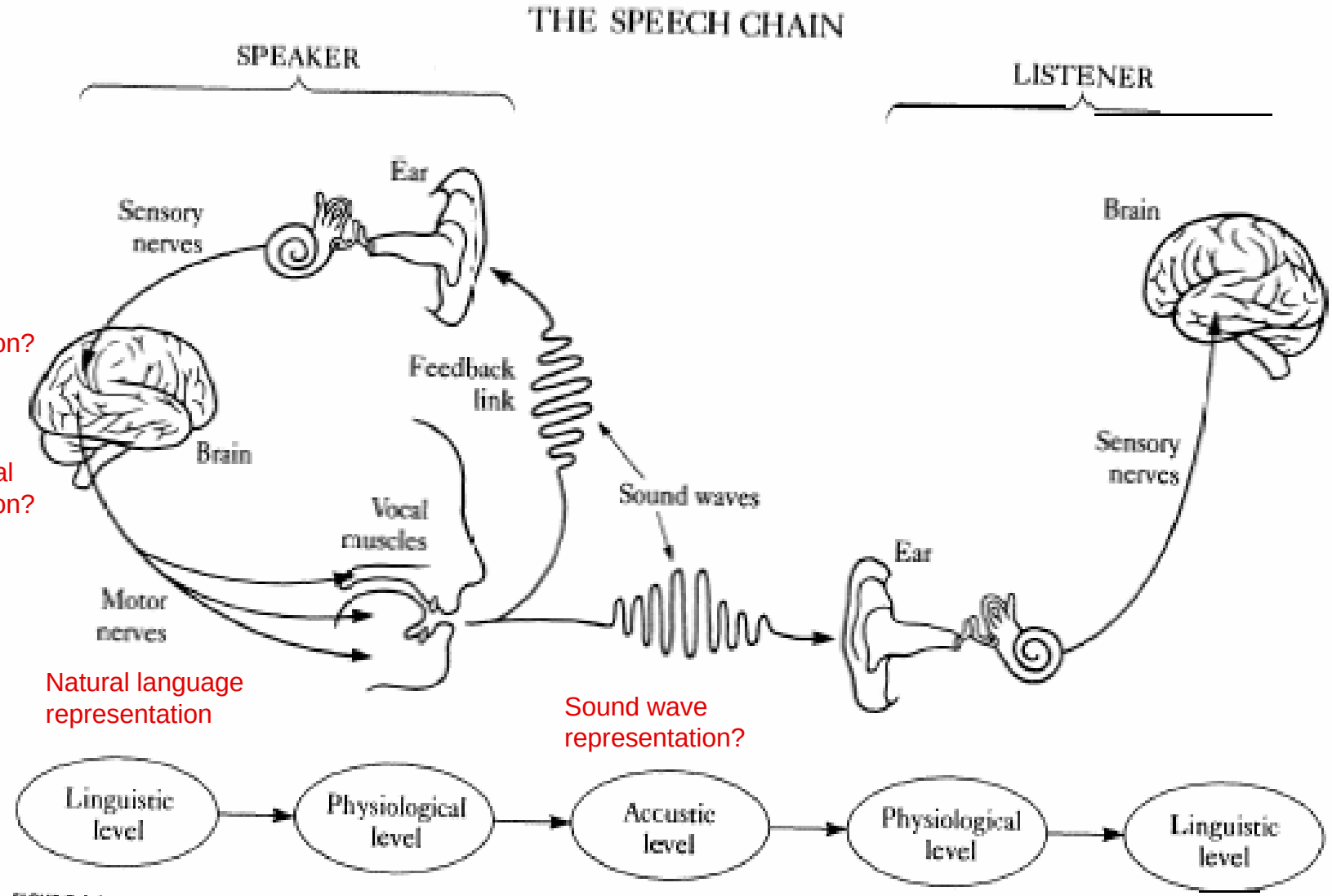


FIGURE 1.1 The speech chain: the different forms of a spoken message in its progress from the brain of the speaker to the brain of the listener.