FAIW

Report

on

The Future of AI Workshop

Date: December 13 - 15, 2002

Venue: Amagi Homestead, IBM Japan

Table of Contents

FUTURE OF AI Workshop

Prospectus

Mankind has witnessed an amazing technological development in the field of Information Technology in the latter half of the 20th century, and during that time, Artificial Intelligence has also grown to be an important and irreplaceable discipline of research and application. Now at the beginning of a new era, evaluating the results of AI research from the past years and providing a sense of direction of where to go for the future should be an effort deeply meaningful to any researcher associated with AI. Especially since AI has begun to attract attention again recently, an important responsibility for us is to investigate the kind of AI technology that will be useful to the society and to propose a guideline that the industry can follow. Forecasting the growth of the new fields of technology such as grid computing and web intelligence should also be an important part of the task in planning out the future.

In the light of the above intension, we have planned to organize a workshop meeting to discuss the future of AI research, calling upon the intellectual ability of the foremost researchers in this field from Japan, the United States, Europe, and elsewhere. Specifically, the meeting will address the following issues:
- What were the important research topics that have been neglected in the past ten years but should be taken up in the near future?
- What will be the important research topics for the next three, five, and ten years?
- What technology will be necessary to have successful AI applications on the Internet? (e.g. web intelligence, knowledge discovery, intelligent human interface, content creation and AI, business grid and its application)

The Workshop meeting will take place on Dec. 14 – 15, 2002 (Reception on Dec. 13), at IBM Japan's Amagi Homestead in Izu.

Following a keynote speech, there will be a series of panel sessions to discuss topics that directly relate to the two main issues of the Workshop stated above. The meeting itself will not be open to public in order to elicit free and active discussion among the participants, but the report may become public through academic and/or commercial publications.

It is with great expectation that we organize the Future of AI Workshop, and we sincerely look forward to the participation of the distinguished AI researchers.

October, 2002

FUTURE OF AI Workshop
Steering Committee Members
Edward A. Feigenbaum, Stanford University
Setsuo Ohsuga, Waseda University
Hiroshi Motoda, Osaka University
Koji Sasaki, AdIn Research, Inc.

FUTURE OF AI Workshop

Outline

1.	Topics of Discussion	a) Currently missing topics in AI research
		b) Challenging research topics in AI for the next 3, 5, 10 years
		c) Necessary technology for AI applications on Internet

2. Date and Place

December 14 – 15, 2002
(Welcome Dinner on December 13)
IBM Japan's Amagi Homestead

3. Meeting Format

2-day closed meeting

4. Language

English

5. Attendees

30 panelists
 -20 from Japan
 -10 from U.S.
Approximately 15 observers from sponsoring organizations

6. Steering Committee

Edward A. Feigenbaum, Stanford University
Setsuo Ohsuga, Waseda University
Hiroshi Motoda, Osaka University
Koji Sasaki, AdIn Research, Inc.

7. Co-sponsors

Air Force Office of Scientific Research
Asian Office of Aerospace Research and Development
Army Research Office-Far East
AdIn Research, Inc.
Alliance Group, Inc.
Canon, Inc.
Fujitsu Limited
Fuji Xerox Co., Ltd.
Hitachi, Ltd.
IBM Japan, Ltd.
NEC Corporation
NTT DoCoMo, Inc.
Advanced BioMedicare Research Inc.

8. Supported by

Ministry of Economy, Trade and Industry
National Science Foundation

9. Coordination

AdIn Research, Inc.

Co-Sponsors:

FUTURE OF AI Workshop **Steering Committee members wish to thank the following for their contribution to the success of this conference:**

Air Force Office of Scientific Research
http://www.afosr.af.mil/

Asian Office of Aerospace Research and Development
http://www.nmjc.org/aoard/

Army Research Office – Far East
http://www.arofe.army.mil/AROindex.htm

*AFOSR/AOARD/ARO-FE support is not intended to imply endorsement by the U.S. federal government.

AdIn Research, Inc.
http://www.adin.co.jp/

Alliance Group, Inc.
http://www.alliance-g.jp/

Canon, Inc.
http://canon.jp

Fujitsu, Limited
http://www.fujitsu.com/

Fuji Xerox Co., Ltd.
http://www.fujixerox.co.jp/

Hitachi, Ltd.
http://www.hitachi.co.jp/

IBM Japan, Ltd.
http://www.ibm.com/jp/

NEC Corporation
http://www.nec.co.jp/

NTT DoCoMo, Inc.
http://www.nttdocomo.co.jp/

Advanced BioMedicare Research Inc.

Supported by:

Ministry of Economy, Trade and Industry
http://www.meti.go.jp/

National Science Foundation
http://www.nsf.gov

<div align="center">

FUTURE OF AI Workshop

Schedule

</div>

December 13, Friday   Registration

Arrival and Registration at IBM Japan's Amagi Homestead by evening

15:00	Courtesy Bus Leaves Ito Station for IBM Amagi
15:40	Bus Arrives IBM Agagi
15:40 - 17:30	Check-in, Break and Unpacking
17:30 - 18:00	Orientation
18:00 - 18:30	Welcome Address / Keynote Speech: Dr. Kazushi Kuse, IBM Tokyo Research Lab.
18:30 - 19:30	*Onsen* (bath) & Changing into *Yukata* (pajama)
19:30 - 20:30	Welcome Dinner
20:30 -	Socializing and Networking

December 14, Saturday   1st Day of Workshop

7:15 - 8:30	Breakfast
8:30 - 8:50	Opening Address: Koji Sasaki, AdIn Research
8:50 - 9:20	Keynote Speech: Prof. Edward Feigenbaum, Stanford University
9:20 - 9:40	Proceed to Study (Seminar) Room
9:40 - 11:10	Session 1: FOUNDATION OF AI
11:10 - 11:30	Break
11:30 - 13:00	Session 2: DISCOVERY
13:00 - 14:30	Lunch
14:30 - 16:00	Session 3: HCI
16:00 - 16:20	Break
16:20 - 18:00	Session 4: AI SYSTEMS
18:00 - 20:00	Dinner
20:00 -	Socializing and Networking

December 15, Sunday   2nd Day of Workshop

7:15 - 9:00	Breakfast
9:00 - 10:30	Session 5: HUMAN-LEVEL INTELLIGENCE
10:30 - 10:50	Break
10:50 - 12:30	Session 6: KNOWLEDGE PROCESSING
12:30 - 14:00	Lunch
14:00 - 15:40	Session 7: SYNTHESIS, SUMMARIES, RESPONSES and other TOPICS
15:40 - 16:00	Closing Remarks
16:00 - 16:30	Break
16:30 - 17:00	Packing and Preparation for Leaving
17:00	Courtesy Bus Leaves IBM Amagi for Ito Station
17:40	Bus Arrives Ito Station

FUTURE OF AI Workshop

List of Panelists

1. FOUNDATIONS OF AI:

- (including) the future of logical knowledge representation and logical reasoning by computer

1. Hiroki Arimura   Kyushu University
2. Stuart Russell   University of California, Berkeley
3. Naonori Ueda   NTT Communication Science Laboratories
4. Akito Sakurai   Keio University *Chair*

2. DISCOVERY:

- machine learning and knowledge discovery, and the future of those research areas

1. Einoshin Suzuki   Yokohama National University
2. Satoru Miyano   University of Tokyo
3. Thomas Dietterich   Oregon State University
4. Hiroshi Motoda   Osaka University *Chair*

3. HCI:

- Human-Computer Interaction and AI, for example in Computer Supported Cooperative Work (CSCW)

1. Yasuyuki Sumi   Advanced Telecommunications Research Institute International
2. Kumiyo Nakakoji   RCAST, University of Tokyo
3. Toru Ishida   Kyoto University
4. Eric Horvitz   Microsoft Research *Chair*

4. AI SYSTEMS:

- scaling up AI systems into large systems such as multi-tasking systems
- possibilities of super-intelligent-systems as an extension of expert system capabilities
- integration of different methods for problem solving

1. Ron Brachman   U.S. Defense Advanced Research Project Agency (=DARPA)
2. Takashi Washio   Osaka University
3. Koji Sasaki   AdIn Research, Inc.
4. Setsuo Ohsuga   Waseda University
5. Edward A. Feigenbaum   Stanford University *Chair*

5. HUMAN-LEVEL INTELLIGENCE:

- computational models of "emotional" processing: are they important?

- possibilities of "human-level intelligence" as an AI vector

- creativity: AI and computational models of creativity

- the importance of coupling the "robotics" work of AI with its "cognitive" work, i.e. "putting a mind in a robot"

1.	Yukio Ohsawa	University of Tsukuba
2.	Masaki Suwa	Chukyo University
3.	Manuela Veloso	Carnegie Mellon University
4.	Naomi Miyake	Chukyo University *Chair*

6. KNOWLEDGE PROCESSING:

- ontologies

- semantic web and intelligent web services

- knowledge management in organizations

1.	Katashi Nagao	Nagoya University
2.	Michael Witbrock	Cycorp, Inc.
3.	Koiti Hasida	National Institute of Advanced Industrial Science and Technology
4.	Ramanathan Guha	IBM Almaden Research Center
5.	Riichiro Mizoguchi	Osaka University *Chair*

7. SYNTHESIS, SUMMARIES, RESPONSES and OTHER TOPICS:

- contemporary definition of AI

- opportunities missed by not coupling closer with neuroscience--to the science of how the human brain works

- topics such as: dissident views; and "what have we missed?"

1.	Koichi Hori	University of Tokyo
2.	Toyoaki Nishida	University of Tokyo *Chair*
3.	Daniel Bobrow	Palo Alto Research Center (=Xerox PARC)
4.	Paul Cohen	University of Massachusetts, Amherst, MA

Attendees Contact Information:
Japanese Panelists
12**(Alphabetical Order)**

NAME	AFFILIATION	E-MAIL
ARIMURA, Hiroki	Kyushu University	arim@i.kyushu-u.ac.jp
HASIDA, Koiti	National Institute of Advanced Industrial Science and Technology (AIST)	hasida.k@aist.go.jp
HORI, Koichi	University of Tokyo	hori@ai.rcast.u-tokyo.ac.jp
ISHIDA, Toru	Kyoto University	ishida@i.kyoto-u.ac.jp
MIYAKE, Naomi	Chukyo University	nmiyake@sccs.chukyo-u.ac.jp
MIYANO, Satoru	University of Tokyo	miyano@ims.u-tokyo.ac.jp
MIZOGUCHI, Riichiro	Osaka University	miz@ei.sanken.osaka-u.ac.jp
MOTODA, Hiroshi	Osaka University	motoda@ar.sanken.osaka-u.ac.jp
NAGAO, Katashi	Nagoya University	nagao@nuie.nagoya-u.ac.jp
NAKAKOJI, Kumiyo	University of Tokyo	kumiyo@kid.rcast.u-tokyo.ac.jp
NISHIDA, Toyoaki	University of Tokyo	nishida@kc.t.u-tokyo.ac.jp
OHSUGA, Setsuo	Waseda University	ohsuga@ohsuga.info.waseda.ac.jp
OHSAWA, Yukio	Tsukuba University	osawa@gssm.otsuka.tsukuba.ac.jp
SAKURAI, Akito	Keio University	sakurai@ae.keio.ac.jp
SASAKI, Koji	AdIn Research	sasaki@adin.co.jp
SUMI, Yasuyuki	Advanced Telecommunications Research Institute International (ATR)	sumi@atr.co.jp
SUWA, Masaki	Chukyo University	suwa@sccs.chukyo-u.ac.jp
SUZUKI, Einoshin	Yokohama National University	suzuki@dnj.ynu.ac.jp
UEDA, Naonori	NTT Communication Science Labs.	ueda@cslab.kecl.ntt.co.jp
WASHIO, Takashi	Osaka University	washio@ar.sanken.osaka-u.ac.jp

Attendees Contact Information:
US Panelists

(Alphabetical Order)

NAME	AFFILIATION	E-MAIL
BOBROW, Daniel	Palo Alto Research Center (Xerox PARC)	bobrow@parc.com
BRACHMAN, Ron	U.S. Defense Advanced Research Projects Agency (DARPA)	rbrachman@darpa.mil
COHEN, Paul	University of Massachusetts, Amherst, MA	cohen@cs.umass.edu
DIETTERICH, Thomas	Oregon State University	tgd@cs.orst.edu
FEIGENBAUM, Edward	Stanford University	feigenbaum@cs.stanford.edu
GUHA, Ramanathan	IBM Almaden Research Center	guha@guha.com
HORVITZ, Eric	Microsoft Research	horvitz@microsoft.com
RUSSELL, Stuart	University of California, Berkeley	russell@cs.berkeley.edu
VELOSO, Manuela	Carnegie Mellon University	veloso@cs.cmu.edu
WITBROCK, Michael	Cycorp, Inc.	witbrock@cyc.com

FAIW

KEYNOTE

Some Scientific and Engineering Challenges for the "Middle" Future Of AI

Edward Feigenbaum

Stanford University

R&D Time Frames for AI

- Near: This Workshop horizon is 5-10 years.
- Far: A 50 year horizon
  - Turing's vision (1950, Turing Test) was 50 year vision (not realized by 2000, but probably will be by 2050)

Envisioning Between Near and Far

- Beyond horizon of this Workshop
- 20-30 year view
  - Maybe sooner with focused efforts
  - (Why do AI scientists have such trouble doing focused efforts, in contrast with physicists, molecular biologists?)

Nature of my Challenge Problems

- Toward the "ultra-intelligent computer" (UIC)
  - Modify Turing's Test appropriately
- Large Knowledge Base for UIC
  - Build a large knowledge base by reading text, reducing knowledge engineering effort by (x 0.1)
  - Distilling from the WWW a huge knowledge base, reducing the cost of knowledge engineering by many orders of magnitude

Turing Test (TT) Reexamined

- What is good about TT?
- Gray's challenge (Turing Award lecture)
- What is wrong (if anything)?
- "Partial Intelligence"
  - "Einstein in a Box"
- Divide-and-conquer strategy for AI science

Various Dimensions of "Partial Intelligence" (Examples)

- Natural Language Understanding
- Computer Vision
- Expert Systems
  - My choice for Divide-and-Conquer
  - On path toward ultra-intelligent computer
  - Early and later examples
  - How limited is the intelligence?

Expert Systems: What Has Been Learned?

- The tens of thousands built can be considered experiments
- What was learned?
  - For an AI system to behave with high levels of performance on complex intellectual tasks, perhaps surpassing human level, it must have extensive knowledge of the domain
  - Knowledge means: terms for entities, descriptions of those entities, relationships that organize the terms and entities for reasoning, symbolic concepts, abstractions, symbolic models of basic processes, fundamental data, a large body of remembered instances, analogies, heuristics for "good guessing," among many other things.

Grand Challenge 1

- Modified Turing Test
- A "Partial" Intelligence
  - "partial" is large, less wide than human experience, but deeper than common sense
  - Uses large domains of science, engineering, biology, or medicine

Challenge 1 continued

- Set up is the same as Turing's
  - Two players and a judge
- One player is an AI System being tested
  - Single- or multi-agent- system
  - Other player is Member of the US National Academy, Royal Academy, Japanese Academy, etc.
  - Judge is also Academy Member

Challenge 1 continued

- N fields of science, engineering, biology, or medicine are chosen
  - I suggest N=3 to make this very difficult test not insurmountably difficult
- If Academy judge can not distinguish AI System from Academy member better than chance level, AI system passes test.
- Passing test in one out of three areas satisfies this Grand Challenge (similar to Gray's 30%).

Grand Challenge 2

- Accomplish the KA for building the knowledge base of the AI System in Challenge 1 by READING TEXT.
- Similar to Reddy's Grand Challenge of "reading a chapter and answering the questions at the end of the chapter."

Grand Challenge 2 continued

- The sequence of "textbooks to read" is itself a Challenge in educational design.
- Human knowledge engineering is allowed, but only to extent of (estimated) 10% of KE necessary to do the whole job manually.
  - To assist inadequate NL understanding
  - To fix incorrect inferences during learning
  - To introduce kernel concepts

Grand Challenge 2 continued

- Success in this Challenge would imply diminishing KE effort for large AI KB systems by one order of magnitude
  - Considered the hallmark of a "revolution."

Grand Challenge 3

- Continuing Education for AI KB Systems or "Keeping up with the literature"
- The AI systems that "won" in Challenge 1 READ the literature of those domains for the next two years; then two more years
- KE is allowed but only to extent of (estimated) 1% of effort for manual KE

Grand Challenge 3 continued

- Repeat the modified Turing Test twice, once after two years, once after four years.
- Challenge 3 is met if the AI System "wins" one of these times
  - I.e. has "kept up with the literature" at least as well as human Academy member.
- Implies reduction in KE effort by two orders of magnitude (super-revolution).

Keynote's Call to Action

- The Japanese and American AI communities together have nearly 10,000 people.
- This Workshop has representatives of the best of these communities
  - Representing many subfields of AI and many diverse viewpoints
- We are here to exchange ideas in an open, direct way
  - Mutual understanding of possibly different viewpoints
  - Clarification of differences
  - Exchange of visions about the future of AI
- Major point of FAIW is <u>discussion</u>, stimulated by panelists

FAIW

SESSIONS

Presentations

and

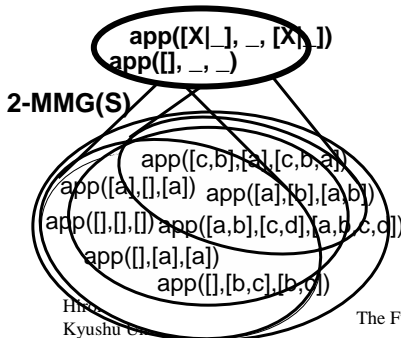Discussions

SESSION1: FOUNDATIONS OF AI

TTT

Hiroki Arimura

Department of Informatics
Kyushu University, Japan

The Future of AI Workshop

December 14 – 15, 2002

Lesson 1. Polynomial-time learning of unions of languages

LGG(S) app(x,y,z)

app([c,b],[a],[c,b,a])
app([a],[],[a]) app([a],[b],[a,b])
app([],[],[])app([a,b],[c,d],[a,b,c,d])
app([],[a],[a])
app([],[b,c],[b,c])

- **Identify** several target languages from the mixture of positive examples alone.

- Known hard; No polynomial time solutions

app([X|_], _, [X|_])
app([], _, _)

2-MMG(S)

app([c,b],[a],[c,b,a])
app([a],[],[a]) app([a],[b],[a,b])
app([],[],[])app([a,b],[c,d],[a,b,c,d])
app([],[a],[a])
app([],[b,c],[b,c])
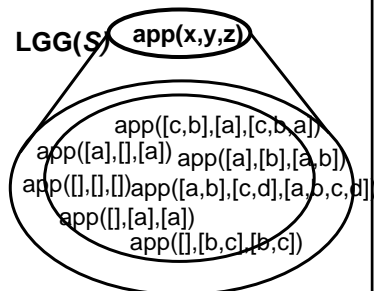
- **We developed** the theory of the minimum multiple generalizations (*k*-**MMG**)

- **Syntactic generalization** for sets of structured objects

- Polynomial time algorithm for computing one of *k*-**MMGs**

- **Sufficient condition**: A combinatorial property on *"a small world"*

27

Slide (top)

Powerset-ordering construction based on subsumption relation

P is more general than Q iff for all *q* in *Q*, there exists some *p* in *P* such that *p* is more general than *q*

P — app([X|_], _, [X|_]) app([], _, _) *p*

q — app([a], _, [a]) app([], a, _) — app([b|_], _, [b|_])

Q

The compactness property:
When the alphabet is **sufficiently large**, the following equality holds

Syntax ⟷ *Semantics*
P is more general than Q ⟷ L(P) is contained in L(Q)

Inductive Inferabilities of the Unions of Pattern Languages

Class of Pattern	C^1	C^2	C^k
General Pattern Language	**Yes** Angluin 79	**Yes** Shinohara 82	
One-variable pattern languages Specific	**Poly** Angluin 79	**Poly** Wright 89	**Poly** Arimura et al. (STACS '94)
	Poly Shinohara 82	**Poly**	**Poly** Arimura et al. STACS '94)('96)
m-variable Regular Pattern Languages	—	**Poly**	**Poly** Arimura et al. (STACS '94)
First-order Terms (Ranked Trees)	**Poly** Reynolds 70 Plotkin 70	**Poly**	**Poly** Arimura et al. (ALT'91)

Solved by our minimum multiple generalization technique

Class of Patterns	Lowerbound	Upperbound
One-variable Patterns (Angluin '80)	k	k+1 (Wright '94)
First-order terms	k	k+1 (Arimura '91)
m-variable Regular Patterns	k+1	2km (Arimura '94)
Regular Patterns (Shinohara '82)	k+1	2k+2 (Arimura '97) k+1 (Sato '98)
Erasing Regular Patterns (Shinohara '83)	k+1	2k+2 (Arimura) k+1 (Sato)

Poly-time learner for unions from positive data = Poly-time MMG + Compactness

The Future of AI Workshop   3

Slide (bottom)

Lesson 2: From Machine Learning to Text Data Mining

Hiroki Arimura (Kyushu University)   Semiar at DCS, Univ. Helsinki, 19 August 2002

Our Research Goal

- Text Mining as an inverse of IR
- Develop **a new access tool** for text data that **Interactively supports human discovery** from large text collections
- **Key: Fast** and **robust** text mining methods

Text Data Mining System

User

Suffix tree

Hiroki Arimura,
Kyushu Univeristy

■ **Basic Idea**

– Reducing **the best *k*-proximity *d*-word association pattern**
– to **the best *d*-dim box** over the rank space

translation by suffix array

...space: ...of all pairs of the suffices of ...ordered lexicographically.

...s and Manber (1990, IPL)

Optimized Rule/Pattern Discovery

Data Mining	Optimized data mining (IBM DM group,1996 - 2000)
Learning Theory	Agnostic PAC learning (1990s)
Statistics	Vapnik Chervonenkis theory (1970s)

Minimization of
- Prediction Error
- Information Entropy
- Gini Index

f(p): impurity function

p: ratio of positives that a pattern matches

Bad rect.
9 positives
9 negatives

Good rect.
8 positives
2 negatives

4

3. Conclusion

- **My experience on Machine Learning and Data Mining**
- **Lesson 1:** Polynomial-time identification of unions of pattern languages from positive data
- **Lesson 2:** Efficient Learning of First-order Horn sentences
- **Lesson 3**: Efficient and Robust Text Mining with Optimal Pattern Discovery
- **Lesson 4**: Web Mining and Information Extraction from Web

- Dichotomies in CS/AI
  - Finite   - Infinite
  - Syntactic   - Semantic
  - Computing   - Learning
  - Deduction   - Induction
  - Concrete   - Abstract

- Details? Principles?
- Tackling the complexity of the world…
  - Logarithm   - Numbers
  - Recursion   - Software
  - Divide&Conquer   - Computation
  - Binary search   - Data
  - Majority Votes   - Learning

SESSION1: Foundations of AI
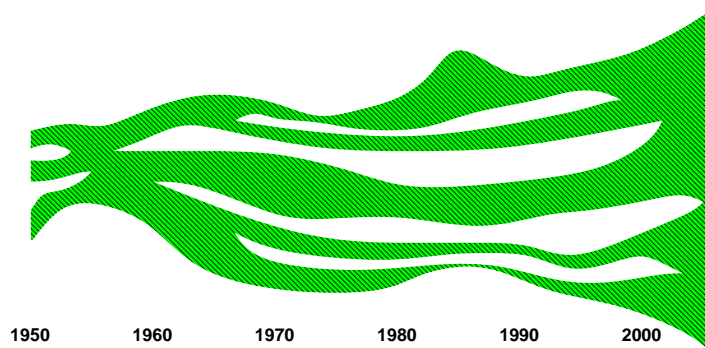
"Key Problems in AI"

Stuart Russell

UC Berkeley

The Future of AI Workshop

December 14 – 15, 2002

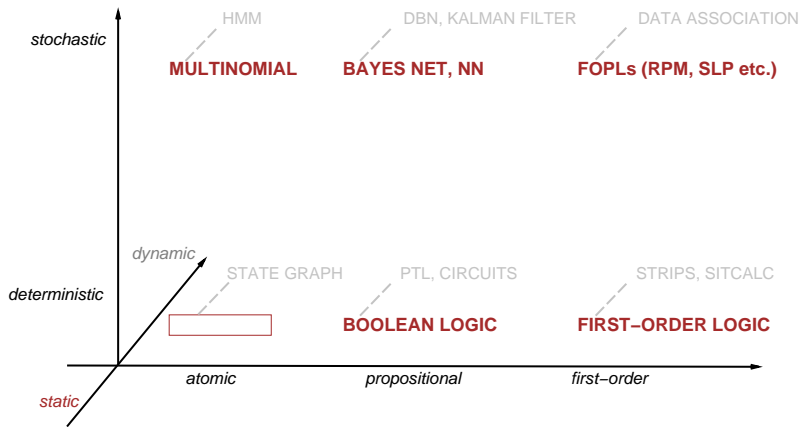## A Brief History of AI

1950   1960   1970   1980   1990   2000

Representation

Representation
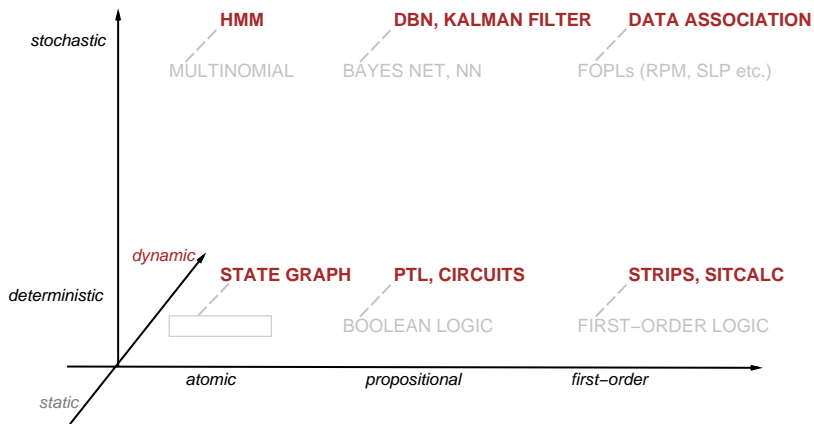
Inference

- Fast SAT algorithms
- Variational approximation algorithms
- MCMC algorithms

Scaling up to 10^6 variables

Decision Making

- Classical planning: new heuristics and algorithms; beginning to understand HTNs
- MDPs and RL: probabilistic state-space search methods
- First-order MDPs (Boutilier et al.) lift *model* but not *plan* representation

Hierarchical Reinforcement Learning
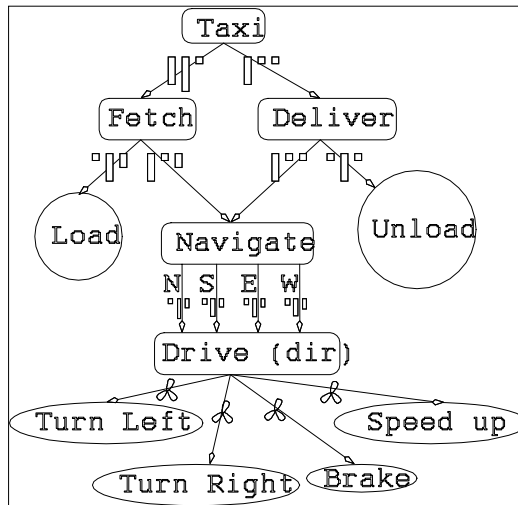
Future Developments

- ***Dynamic open-world*** FOPL agents
- Full integration of KR with vision and NLP
- Hierarchical metalevel reinforcement learning $\rightarrow$ bounded optimality
- EBL and universal subgoaling in a decision-theoretic context
- Closed-loop cumulative learning

Session 1: Foundation of AI

Developing Mathematical Methods for Solving Symbol and Pattern Fusion

Naonori Ueda
**NTT Communication Science Labs.,
Kyoto Japan**

Why did conventional studies fail?

Mainly lack of mathematical formalization.

Almost were ad hoc methods
(ex. Image understanding)

Creating a new concept is necessary, but it itself is insufficient.

Creating a new method for realizing the concept.

How to approach is essentially independent of a target problem.
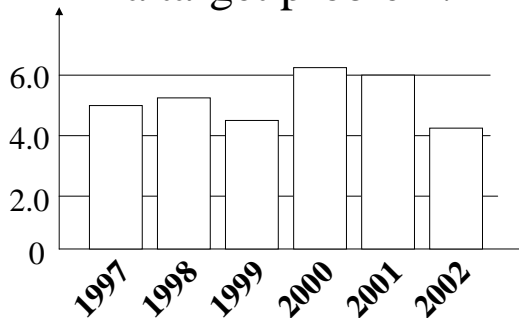
# of equations/paper in Japanese AI Journal

⬇

# of mathematical papers was very small !

What should we do?

One key candidate: Web analysis

Characteristics of the Web

Multimedia data analysis (text, image, speech, etc)

- high-dimensional, sparse data
- huge data

Social data analysis (community structure)

- complex link analysis
- clustering

How should we do?

Importance of Semantic Mapping

Data $\longrightarrow$ **concept**

$\uparrow$ **Semantic correspondence**

$\downarrow$

Mathematical model
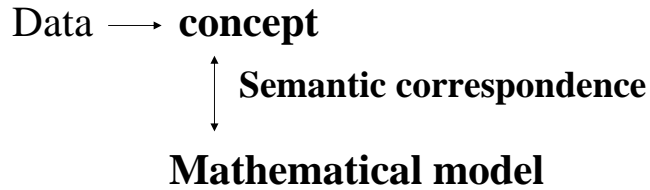
Example 1: Hyperlink analysis

(Kleinberg, 1999)

Concept: hub & authority degrees

hubs ⟶ authorities
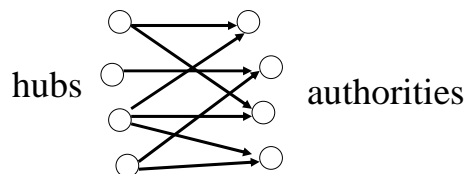
Formulation: $a^{(t+1)} = A^T h^{(t)} = (A^T A) a^{(t)}$

$h^{(t+1)} = A a^{(t+1)} = (A A^T) h^{(t)}$

Hub (authority) degree can be obtained by solving eigen vectors.

Example 2: Latent Semantic Analysis (LSA)

(Deerwester et al., 1990)

Concept: Vector space model of
Words & Documents

Formulation: SVD of a word-to-document association matrix.

$$w_1, w_2, \ldots, w_V \begin{array}{c} d_1 \; d_2 \cdots d_N \\ \boxed{X} \end{array} \xrightarrow{SVD} \underset{U_k}{[\boldsymbol{u}_1 \cdots \boldsymbol{u}_k]} \Lambda_k \underset{V_k^T}{[\boldsymbol{v}_1 \cdots \boldsymbol{v}_k]^T}$$

Document sub space   **Word subspace**
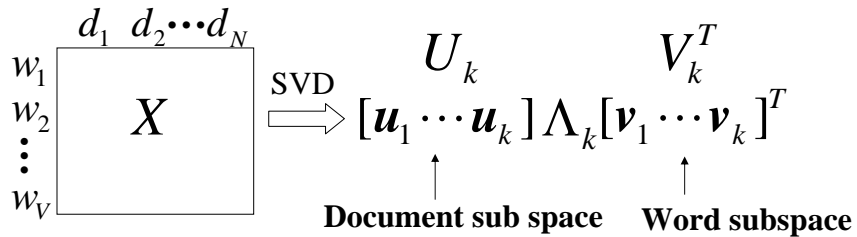
Example 3: Probabilistic Model for
Multi-labeled Text   (Ueda & Saito, 2002)

Concept: Generative model for multi-labeled text

Formulation: Parametric Mixture of Multinomial distribution

> Each parameter vector corresponding to multi-topic class
> can be represented by **a linear combination of *L* basis vectors**

L=3 case $\boxed{\text{math}\,\Theta_1}$ $\boxed{\text{sports}\;\Theta_2}$ $\boxed{\text{music}\;\Theta_3}$

math & sports $\Theta_{1,2}$ math & music $\Theta_{1,3}$ sports & music $\Theta_{2,3}$

math & sports & music $\Theta_{1,2,3}$

$\Theta_{1,2} = c_1\Theta_1 + c_2\Theta_2$

$\Theta_{1,2,3} = c_1\Theta_1 + c_2\Theta_2 + c_3\Theta_3$
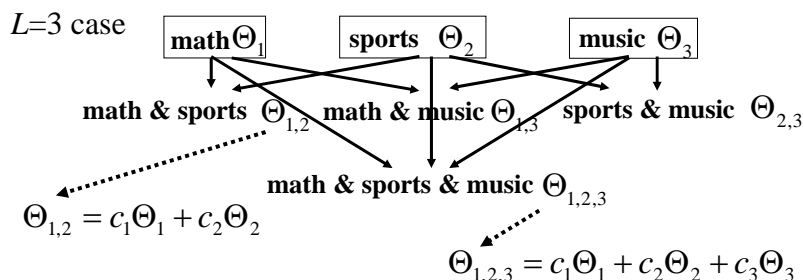
SESSION1: Foundation of AI

"From Static to Dynamical AI"

Akito Sakurai

Keio University

The Future of AI Workshop

December 14 – 15, 2002

Static vs. Dynamical

Static AI = algorithm-based AI:

    Tightly coupled with Turing paradigm where
        a computation is done within a computing mechanism,
        algorithmic time step is independent of real clock, and
        the time-line used in AI is an artificial one-dimensional sequence.
    Successful for many small and large applications, but
    It has been more similar to algorithms than to life.
    If new applications were expected, new approach is required.
    A trial is to focus on interactions, i.e., an action is evoked and
        clock is advanced step by step by extraneous stimuli.

Dynamical AI = interactive and temporal computation-
    based AI:

    Interaction is more focused on than algorithm is.
    Temporal representation is inevitable
        forced by introduction of advanced devices.

Architectural Impact

Quantum computer in some form might be used,

"Real" one is hard, but, say, molecular arrays are realistic,

which is no more deterministic nor reliable

Could be made reliable enough as a discrete logic unit by aggregating many, but it would deprive of its advantages.

which has no long connections nor clock lines.

Different from the brain (long connections exist)

Same as the brain (no central/global clocks)
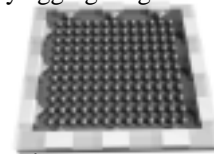
Symbolic processing capability is to be built on it,

©T.Ogawa, Ehime Univ.

as in the human brain, to support the algorithmic faculty.

AI	logic and algorithms
algorithmic basis	AI (natural language)
discrete logic	quantum devices
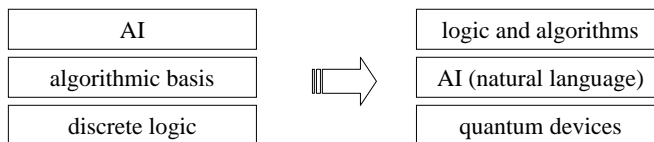
Temporal Representation

Hints from the brain, again

We have already learned from intelligence and the brain

Turing machine, machine learning, multi-agent,,,

Threshold logic, neural networks,,,,

Temporal representations are abundant in cortex

correlation of spikes of neurons,

synfire chain, etc.

There remains a problem if temporal rep. is essential

Is it just another dimension that can be subsumed into graph rep.?

But if we have unreliable distributed processing units, local connections, and processing time limitation, we have to utilize temporal coding/processing.

Interaction vs. algorithm

Interaction common in AI (Simon 1969, Brooks 1991):

Behavioral complexity might reflects environmental complexity.

Interaction in computation (Wegner 1997, Leeuwen 2000,2001):

An Interactive TM with advice is more powerful than a TM

(where "upgrade"(=advice) is important).

AI by interactive TMs with advice(?) (not formulated yet)

May give us new paradigms of computation,

It might not change computability, though.

since areas in the brain might be considered as an approximation to infinitely many ITMs with advice.

Summary

AI, in future,

is a basis of algorithms and computations,

bridging between discrete logic (over the AI) and unreliable or stochastic but far compact and fast hardware.

<= intelligence of low to high abstraction level

where the data-structure is spatio-temporal,

<= "temporal coding"

interacts with outer world, and

<= "interactive computation"

exhibits life-like intelligence

T.M.	AI		mathematics/logic
		F.AI	logical thinking
			natural language
			primitive brain
	AI	F.AI	neural networks

Session 1   Discussion

Paul Cohen:
When I saw that this was going to be a session on foundations, I thought I might hear some more philosophy. A philosopher asks what's a mental state? How do mental states have meaning? How does the meaning of the mental state explain behavior or predict behavior? How do meanings come to be associated or learned? What kind of representations do you need to carry particular kinds of meanings? If you're going to have compositional representations how do you have composition of meanings? How do you handle particular kinds of knowledge about the world, in particular knowledge about time in your presentation, knowledge about objects in Stuart's presentation? I thought we were going to hear more about the philosophical underpinnings of the work that's going on in representation, and I'm not sure whether the absence of that discussion indicates that everybody knows what the philosophical foundations are and they're implicit or nobody knows what the philosophical foundations are and we're struggling.

Stuart Russell:
I think you have to say that people are building representations successfully, and once the technical foundations of those are clear then philosophical issues have in some sense evaporated. So the question is, take one of your favorite representations and say, "What philosophical issues remain unresolved?"

Paul Cohen:
Well, for example, how would you realize the notion of a fluent in the situation calculus? Are you satisfied that everything that was intended by the notion of a fluent in the situation calculus is in fact realized by your favorite representation? Have we in fact got this idea right of a state which, although it may be changing, is not changing in any fundamental way with respect to the behavior of the agent? Are we comfortable that we can explain how you can get from a time series of sensory data to symbolic representations that carry meanings that effect the behavior of the agent? I know we can build things. My question is, do those things actually realize any philosophical position, or have we simply said, "Look, the philosophy is irrelevant. We're building things, and that's enough?"

Stuart Russell:
I think if the systems that you build don't break then you're OK. What I was arguing in my talk was that if you were to build—let's say you have dynamic Bayes nets which are capable of interpreting sensor data over time and constructing representations of the state. But they break because they don't represent objects and relations. They can't scale up and they can't produce symbolic knowledge of an interesting kind, whereas if you look at situation calculus it doesn't handle uncertainties. You can't connect it to the real world, and so, again, they break. So instead of saying—maybe we are saying the same thing, that the philosophical issue cashes out in the sense that the system breaks and you need to figure out how to get around this limitation.

Daniel Bobrow:
You know, I think it's different in that Grade actually talks about a correspondence between two different systems, one of which is your system of expectations and the other of which is the kind of behaviors that you are seeing in the system. The fact that you don't see fluence

appropriately in situation calculus, or you don't see objects in Bayes net has to do with the fact that you, Stuart Russell, want to have objects because you would like to have something that corresponds to your notion of objects.

Stuart Russell:
No, really, not at all. It's because I want, for example, to represent the rules of chess. And to do that in a propositional language requires about a million pages of rules. That's totally out of the question. It's not a question of what I think it ought to be. It's just not practical.

Daniel Bobrow:
So it's a question of comparing that representation with some other representation, and some notions of your expectations. And where do these expectations arise from? I think that's where this philosophical stance comes in. What's wrong with a million of these steps?

Manuela Veloso:
I have a question for Stuart and also for these other talks. When we build a system and we try to solve a very concrete problem, I never feel the lack of first order predicate logic, probabilistic language. I would feel the lack, if that is actually what drives my inability to create a full artificial intelligent being. So it's interesting that I find bottlenecks at a completely different level. I cannot handle infinite objects, I cannot represent unboundedness, and I cannot think about universal quantification at the probabilistic level. So my struggles with thinking about artificial intelligence at the system level involve a lot of non-linearities. The fact that there are switches, the fact that there are transitions that have to be addressed with different algorithms, the fact that there is a multitude of models and hypotheses. It's not the lack of first order probabilistic logic. So what I find most striking from a foundational point of view, even with logic and all of that, is you need a mathematical framework that allows me to switch, that allows me to jump from one place to the other in a principled way.

Yukio Osawa:
Yes. I think her comment is hitting the point. However, it is very difficult to find a chance to switch to a completely different context from where we are in because we have many candidates of the next context to switch. We have to find a very deep level understanding of the environment for judging which context to switch and to which context we can, and we should, and we will switch. The emotional and philosophical and various other things are involved in the productive judgment. Currently I am doing studies on rare events in business applications. And in these studies I am feeling very strongly that it cannot be solved completely by computers alone. The more important thing is to make interaction between the environment, and humans, and the computer. I think the interaction with the outer world is the difficult key for the future of AI.

Takashi Washio:
I think we need comments from a philosophical view. Thinking about switching from a framework of AI to another framework of AI depending on the situation and system status, at least I think that both methods have to share some background knowledge. Otherwise they cannot switch and still look smart in doing so. Maybe for the future of AI we have to work on how to share the knowledge, or axioms, or premises among different frameworks like Bayes nets and others. And we also need to share such knowledge with humans.

Yukio Osawa:
But also we have to share knowledge with the environment. Much knowledge exists in the external world outside of the humans and outside of computers.

Manuela Veloso:
Yes, but now from a foundational point of view, I keep thinking about where are we studying changes of representation, where are we studying suitability of particular representations for a particular context? How are we studying representations that, at a certain point should be like this but need to switch? So where is the science of switching in AI? How do we switch? What's the glue? How are we going to look at things sometimes as real logic and real inference, and some other times as really tossing a coin, randomized choices? Well-defined inference to complete random decisions, and how do we actually coordinate this in the AI system and on top of that combine it with the outside world? One thing I like to see in the foundation of the future is this switching theory.

Eric Horvitz:
On the relevance of philosophical foundations to building effective systems, we should loft the question: To what extent are there key conceptual bottlenecks in the realm of our understanding of the philosophical foundations of intelligence that are genuinely limiting our progress in building useful systems? Where are the hot spots here, and how should we focus our attention on key philosophical issues?

Manuela Veloso:
I'm trying to address the question of philosophical foundations also by seeing that philosophically, the realm of intelligence has to do with switching, with the change, with the flexibility.

Thomas Dietterich:
But maybe it's just because all the representations you are using are inadequate and if Stuart can give us these unified representations, then the switches would be unnecessary. It would just fall right out of the realm.

Manuela Veloso:
That is actually something that I claim is not true. I don't claim that we are not lacking some representation but I have a hard time claiming that we're going to find a universally omnipotent representation. I find it hard to accept that.

Yukio Osawa:
Humans switch the representation when they can deal with analogies in their mind. Humans have rich experience so they can connect the current experience with some other past experience.

Ramanathan Guha:
Stuart and you are not saying different things. Using McCarthy's terminology, he is looking at the entire probable and epistemological point of view, which is multiple representations, multiple things and we need an overarching theory of how all these things can be put together. And you are looking at it from the puristic but processing point of view, saying that there are different tasks, which require different representations and you need to do the switching and so on and so forth. And we need both. The existence, me for one does not…

Manuela Veloso:
I didn't understand these as multiple representations.

Stuart Russell:
I certainly was not saying that dynamic open world first order probabilistic languages are the only bottleneck. I listed six areas for future work and that was just one of them. And specifically, I was saying that we need that type of representation in order to construct real large symbolic open spaces from raw data. Now, that is not something that you're currently working on and so you don't see that as a bottleneck. But I think if you want a robot like the ones that you're working on for Robocup, and if each of the robots had individualized behavior and they could bring on substitutes, you would need to recognize that, "Oh, this is John the robot who's coming on."

Manuela Veloso:
The challenge is actually by observation to figure out that this is the strong one not the stronger or whatever. The challenge is at a completely different level.

SESSION2: DISCOVERY

"Data Collection and Preprocessing
Face Reality"

Einoshin Suzuki

Yokohama National University

The Future of AI Workshop

December 14 – 15, 2002

---

1. Background

Analysis &evaluation

Pattern extraction

Transformation

Preprocessing

Selection

data

Target
data

Processed
data

Transformed data

Pattern

Useful
know
-ledge

Data mining
process

[Fayyad 1996]
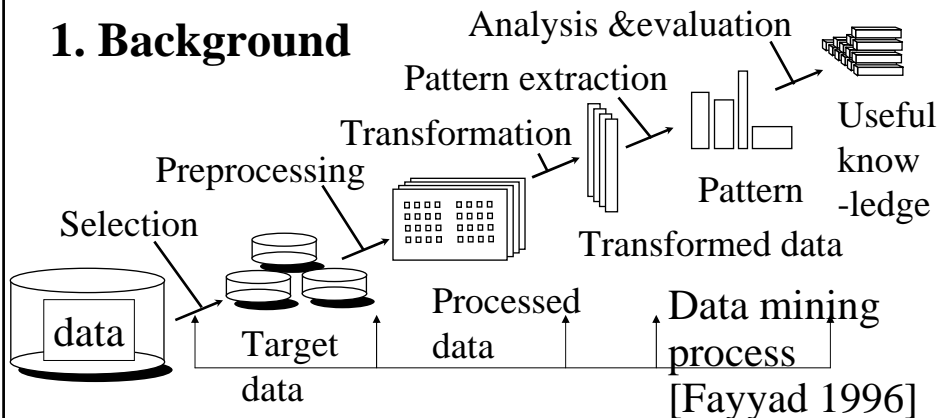
*Problems: Burden of data collection and
data preprocessing*

*e.g.) 800 hours spent for organizing KDD Cup
2000 [Kohavi SIGKDD Explorations 2001]*

2. Old Media as Data Sources

New media ⟹ *Accessibility to contents*
(WWW, DB, digitized archives, etc.)

DB   **archives**
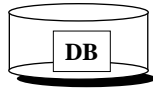
Old media ⟹ *Wealth of contents*
(TV, newspapers, magazines, radio, books, etc.)

RADIO

Next 10 years: Nontrivial discovery from stream
  data collected from old media

3. Quality of Structured Data

Structured data ⟹ *important aspect of reality*
(various kinds of graphs, time-series, etc.)

Heterogeneous structured data ⟹ *real problems*
(e.g. medical records of GOT for patients)
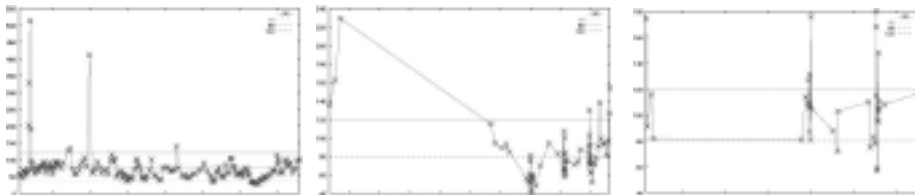
Next 10 years: Highly-automated discovery from
  massive heterogeneous structured data

4. Limitation of Current Technologies

1. Data Collection

Pattern recognition for images [IEEE Trans. PAMI]

⟹ *Data collection itself is its objectives*

Research on WWW [Proc. WWW 1994-2002]

⟹ *Fewer contents than in traditional media*

2. Quality in Structured Data

Graph mining [Dehaspe 1998, Motoda group 2000-],

Time-series Mining [Berndt 1995, Keogh 1997-]

⟹ *Need preprocessing to certificate qualified mining for highly heterogeneous data*

---

5. Conclusion and First Steps

1. Old Media as Data Sources: requires nontrivial but feasible objectives

⟹ *Plan to begin publishing in 2 years*

2. Quality of Structured Data: Medical examination data can be a promising source

⟹ *Expect to begin publishing from 2003*

Are these topics novel?

⟹ *Coming research should face reality (problem settings, scale, robustness, etc.)*

Session 2 Panel 1   Discussions

Thomas Dietterich:
How much do you think the problem with pre-processing comes from the fact that the data is structured but our learning algorithms don't work well with structured data?

Einoshin Suzuki:
I would say half comes from pre-processed data. But this is our view in the early 90's. Of course, you'll know that these similar steps have been coupled but they are divided because the data must come in heterogeneously. But when I read recent papers, I see more than one algorithm, which deals with structure directly and what we are working on is such kind of algorithms.

Hiroshi Motoda:
I guess it means even if you have a very nice algorithm that can handle structural data, you still need knowledge to do a good pre-processing.

SESSION2: DISCOVERY

"Life *in Silico*"

Satoru Miyano

University of Tokyo

The Future of AI Workshop

December 14 – 15, 2002

---

1. Stragety for Life *in Silico*

Understanding of Life as System
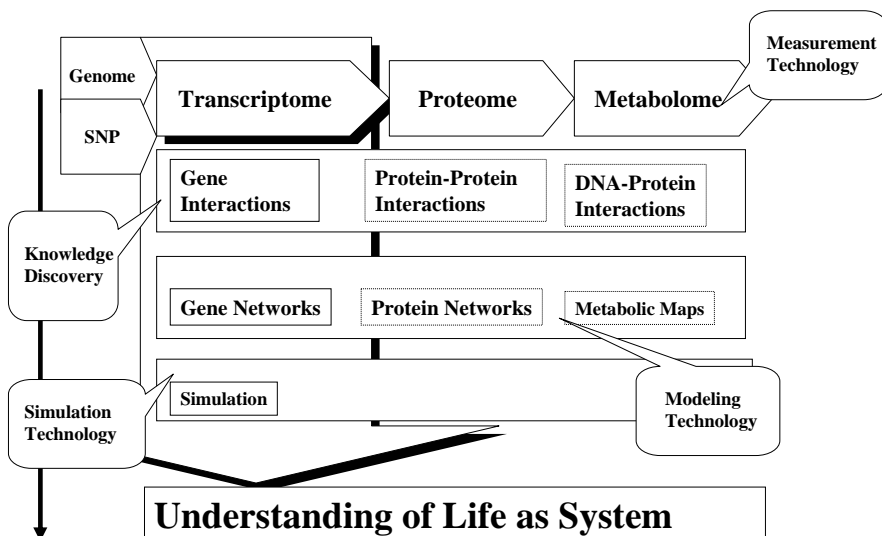
2. Mining Network Information from Genome-Wide Heterogeneous Biological Data

DNA Sequences
Amino Acid Sequences
Protein Structures
Gene Expression Data
  cDNA Microarrays
  DNA Chips
Protein-Protein/DNA
  Interactions
  Protein Chips
  MALDI/TOF MS
Literature
 MEDLINE database

(C)2002 Satoru Miyano
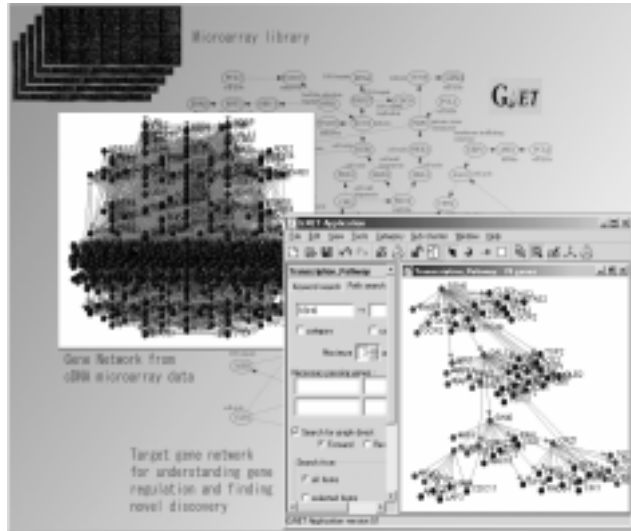
3. Modeling, Simulation and Discovery

Modeling and
Simulation of
Intra-Celluar and
Inter-Cellular
Systems
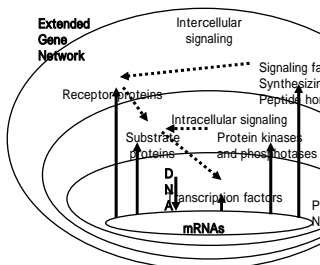
(C)2002 Satoru Miyano
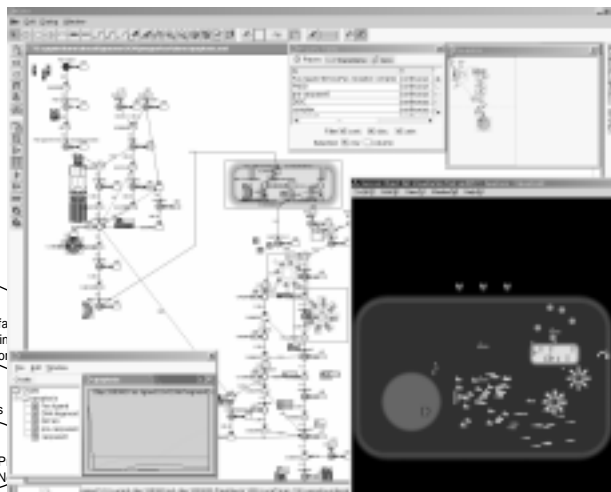
4. Conclusion

- 21st Century is *Life Science Century*
- Life *in Silico*
- Mining Network Information from *Heterogenous* Biological Data
- Modeling, Simulation and *Discovery*
- *Emergence of New Science for Understanding Life*
    - Fusion of Biology, AI, CS, Physics, etc.

Session 2 Panel 2 Discussions

Edward Feigenbaum:
Is there much activity in Japan about "The fusion of biology, AI, CS, physics," or are you relatively unique?

Satoru Miyano:
Relatively unique. But in the United States, Leroy Hood at the Institute for Systems Biology has created systems biology by combining various kinds of fields, computer science, AI, biology, physics, chemistry, and so on in order to understand the life as a system.

Edward Feigenbaum:
In view of the morning's discussion, I thought it very interesting that you reflected on the idea that we consider biology a very powerful science. It achieves great things, but it doesn't do it with mathematics. In fact, they go out of their way to say they don't do it with mathematics, they don't want to know the mathematics. At various times Doug Lenat and I talked about "physics envy." Physics envy means you love mathematics; you think mathematics is what you have to have underneath your science. But that's not true.

Daniel Bobrow:
What is the form of a result in biology such that you can make it productive? You say you have results, you show simulations. What is the underlying representation of the simulation, of the processes that are happening?

Satoru Miyano:
Such networks are described using XML documents. And the dynamics are written with Petri nets. However, we need some GUI so that biologists need not take care of the details of such mathematics.

Hiroshi Motoda:
Isn't it something that the system learned, that human biologists that are critical about using this language want to know?

Satoru Miyano:
Yes. So, we appreciate the biologist's intuition and knowledge very much. Biological knowledge is very hard to encode.

Stuart Russell:
I do a little bit of work. Sir Roger Brent at the Molecular Sciences Institute also has a very large project going with this kind of systems biology. And it seems to me from discussions with the biologists that they do have a mathematical framework. Namely, what they do is to write down lots and lots of chemical reactions that mathematically describe these processes. And then in many cases people are doing quantitative simulations of those because in the cell, particularly when you have binding, the numbers of molecules involved are one, or two, or eight. You can't use stochastic differential equations. You actually have to use a discrete stochastic simulation in order to see how the system works. So, it's not clear to me that biologists have no mathematical

framework, it just doesn't look like the same type of mathematics that other fields use. The reason for that is because the phenomena are slightly different.

Daniel Bobrow:
We heard that mathematical equations don't count logic, and we hear that they don't use mathematics because they use discrete mathematics. I think it's the question of formal representation versus whether it's called mathematics is something that we should not buy into, of all people. When John McCarthy was doing LISP, he was doing the original symbolic reasoning with mathematics. It was based on Kleene work, and there's a long history of these things. So, we should be careful in talking about the fact that they don't use mathematics.

Edward Feigenbaum:
That is a misreading of history. List processing was not invented by McCarthy; it was invented by Newell and Simon out of computer science ideas.

Daniel Bobrow:
I said the invention of LISP; I didn't say the invention of list processing. All I am saying is that there is a lot of mathematics that's being used and a lot of what we do in terms of our algorithmic development and our processing of logic and even rule-based systems are moment discrete mathematics. And so I am just saying that by saying we don't have mathematics, we're putting ourselves in a spot, which I don't think is appropriate.

Manuela Veloso:
So, what is mathematics? I really am genuinely asking.

Satoru Miyano:
Mathematics is something being taught at the department of mathematics.

Manuela Veloso:
Do they teach rule-based systems in the department of mathematics? What I am trying to say is this: Are we going to subscribe to a principle where there are no boundaries? That's fine. We can say everything is mathematics. I think that Ed's comment was being like, "Okay, there is another eventual underlying assumption of what mathematics is." So, I don't know. I am genuinely asking. Is there a concept of what is mathematics or not?

Stuart Russell:
There needs to be some way of stating things in a way that is officially precise that you can put two things together and derive a third and be confident of the outcome. So, when biologists write down these chemical pathway descriptions in the form of equations, they can say, "Well, look. If this equation and this equation exist, then there is a pathway from one end to the other and I confirm that." And when they draw big graphs and grids and nodes, the equations are linked. It's like saying that this species is produced from this one and this one and here is the pathway. That is a mathematical representation. It wouldn't work if they didn't have a precise notion of what the basic components mean.

Michael Witbrock:
When necessary, they use real mathematics too. They use reaction diffusion equations, which are differential equations to describe how growth hormones are distributed in the structure of your brain. So, I don't think biology is free of equations; chemical equations are sort of equations, mathematical logic is another kind of equations…

Takashi Washio:
I think a similar discussion applies to physics also. I think in physics, in the elementary particle level, they do not use the ordinary notion of mathematics even if the wave function is represented by the Schroedinger equation. So, I think it's a matter of the level of description.

Eric Horvitz:
I'd like to comment that whether the base-level mathematics are qualitative or more detailed quantitative models--such as numerically intensive linear programming or differential equations--the models are used to reason formally or informally about the likelihood of theories, and about such relationships as causation. No matter how the base models are constructed, we wish to reason about the predictive power of those models, for example to assign belief to competing theories for what's going on in a cell. Reasoning methods within or on top of the base models can be used to assist with validating our understanding of the fundamental causal structure of phenomena. We're starting to see in the UAI [Uncertainty in Artificial Intelligence] and mainstream AI communities the application of mathematics and theories for identifying causal influences, based on probabilistic and constraint-based reasoning. Tools for assisting with scientific theory confirmation may one day be used commonly in the course of discussing alternative explanations for phenomena, for planning experimental programs, and for making decisions about funding different directions in scientific research.

Thomas Dietterich:
I sometimes kid my wife who is a molecular geneticist that biology is like a giant community of automobile mechanics and they are all studying the same thing, which is a 1952 Mercedes Benz, and they publish papers of the form, "Oh, this cable goes from the brake to the brake pedal." And that's one paper and so on. And the genome project is essentially trying to build a map of a car. Now, the fundamental biological question is why is this here rather than something else? And how does it all work as a system? And they are just beginning to get to that stage so they are in some sense, in a pre-theoretical phase where they're collecting data and trying to understand these sub-systems and trying to compare them across, Mercedes being compared to Fords and Toyotas, to see what are the design principles occurring across all of these viewpoints. So, physicists don't do this kind of thing where they might publish, "Well, I found the locations of all these molecules in here and these objects over here." That was done in the early days of astronomy when people were recording the locus of planets and so forth. And that then initiated the search for the principles. So, I think the difficulty biologists face is that the phenomena are so much more complex than the ones in physics that it is going to take a long time to get to the reductionsim phase.

SESSION 2: DISCOVERY

"Learning and Prior Knowledge"

Thomas G. Dietterich

Oregon State University

The Future of AI Workshop

December 14 – 15, 2002

Our Dream

Prior Knowledge

↓

Learning ← Raw Data and Experiment Results

↓

Knowledge Revised and Extended
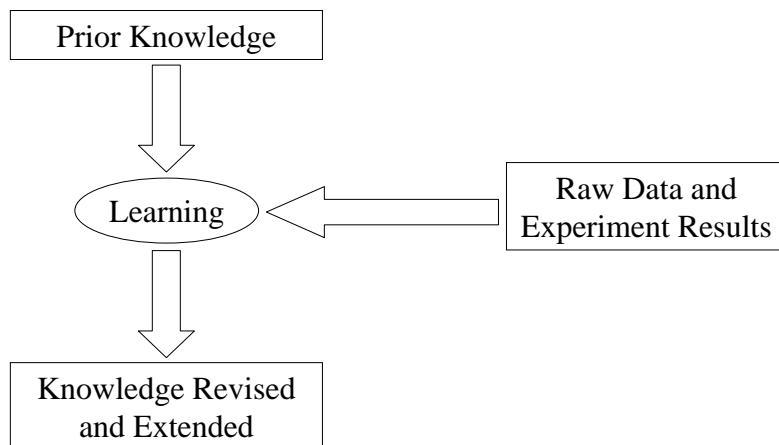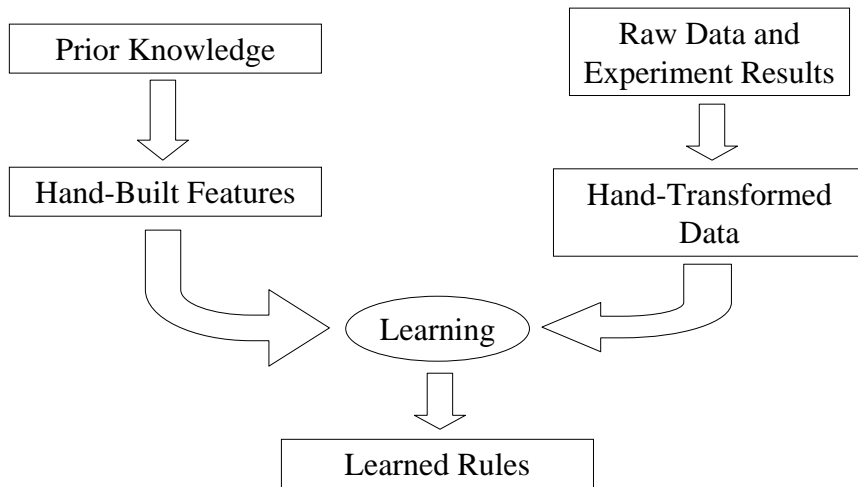
Thomas G. Dietterich   The Future of AI Workshop   2

Current Practice

Prior Knowledge		Raw Data and Experiment Results

Hand-Built Features → → Learning ← ← Hand-Transformed Data

Learning → Learned Rules

Why is this a problem?

- Learned knowledge cannot be combined with prior knowledge to support inference
- Therefore, knowledge cannot accumulate
- Each machine learning application requires designing the features and transforming the data. Labor intensive!

Why do we do it this way?

- To represent our prior knowledge (in general) we need a very expressive knowledge representation
  - e.g., CYC, Classic, differential eqns
- Statistical learning is not feasible in these representations:
  - Statistical problem: too rich
  - Computational problem: too complex to search and optimize

Research Challenges

- Automating the design of features
- Automating the transformation of the data
- Understanding how to update and extend the prior knowledge to incorporate the learned rules
  - How do we represent evidential relationships and uncertainties in the original ontology?

Conclusion

- Current ML systems required hand-constructed features and manually-transformed training data
- There are good computational and statistical reasons for this
- We need to automate the design of features and the transformation of the data
- We need to understand how to update our knowledge to incorporate learned regularities.

Session 2 Panel 3   Discussions

Daniel Bobrow:
One puzzle for me, always, in learning has to do with its disconnection from any action, that you talk about learning and you talk about generating, effectively, formula. But I think that when you think about the learning, where it comes from the human, is the need to actually be effective in the world, and in the feedback in the world about what has to happen. And so maybe you could say something about how you connect this back.

Thomas Dietterich:
Actually, one of the big controversies in machine learning community is between the decision-oriented folks and the model-oriented folks. In the decision-oriented approach, if you think about something like character recognition, we think that in learning we want to just learn a function and a match from an image to a decision about what characters are in that image. So, stated purely in terms of action, we might have a neural network or a support vector machine or a decision tree that just tries to make that mapping correctly.

Then there's the modeling approach that says, "Let's first learn a model of how the characters are created." So, you might say we have a Bayesian network of some kind. There is a very elaborate one of such kind done by Jeff Hinten's group in which they imagine there is a sort of platonic integer 5, and then there's a space of possible ways it could be distorted. And then there's a process that generates the ink that goes under the paper. And that's how the 5 is generated. And so we build a model of that and then you have, given that particular observed 5, you basically have to apply Bayes theorem to invert that model and then make a maximum likelihood or maximum expected utility decision about it. So, that's how you connect the learned models to the decision-making.

That's sort of in the one shot world. Then there's all the world of reinforcement learning in which there is a similar kind of dichotomy except it gets more complicated because you have sequences of decisions over time. But one family of methods there, the so-called model based ones that first have to learn a model of the world, now has a discrete time Markovian system or dynamic phasing network or something like that. And then, once you have that model, and assuming you also have the utility function, the cause function, the work function or whatever it is you're trying to optimize, then compute an optimal decision-making policy. So, that's the model-based approach.

Then there's the model-free approach that just says, "Well, let's consider the space of all possible policies maybe representatives, computer programs or some sort of prioritized policies. And let's search directly in that space." And right now, the pendulum is flung to that end. So, machine learning is very much connected to action.

Daniel Bobrow:
But when you talked about these problems, you talked about the need for doing this transformation of data. So, connect that problem to what I am talking about to what you just said.

Thomas Dietterich:

So, when I'm building a system for detecting fraud in credit card transactions, first I have to decide what kinds of features I'm going to collect out of the transaction history for a particular customer up to that moment. And that's where the hard feature of engineering happens because potentially, there are all kinds of features in that history that might be useful – the locations, the amounts, the names of the companies, how many miles between this transaction and the previous one, does this person move around a lot? And so, I use lots of prior knowledge about human behavior to design those features. Then I run the learning algorithm and I get out a rule that makes decisions. And actually, in this case, I just implement that rule. So, we have things like the Nielsen company selling all those neural networks for this kind of thing and they just make that decision. But what I don't ever do is close the book and go back and say, "Well, was my prior knowledge about human behavior actually correct?" And maybe if it was or maybe if it wasn't, if I could improve that knowledge, maybe I could reduce the false positive rates in these systems or reduce the false negative rates. But we don't seem to be able to close that loop so the systems never become any more accurate than they were in that first iteration.

Daniel Bobrow:

Could you not go back and try to automatically extract features, which would make a difference between the actions you wish you had taken and the actions you did take?

Thomas Dietterich:

Yes. Or you could ask also what other kinds of data could I collect that would help me out in this analysis.

Einoshin Suzuki:

Could you give us your opinion of constructive induction?

Thomas Dietterich:

So, constructive induction, most of the work there is assuming that I start with some raw features and then I have some feature combinations techniques. And I think there are a lot of cognizant ideas there but most of it is not guided by knowledge. I was working for instance on a project where I was trying to predict grasshopper infestations and one of the things that we knew about the grasshopper life cycle is that they grow as eggs in the soil until a certain date when they hatch, and then they come out of the soil and they start eating things. And that's when the farmers start to get angry. And the features that are relevant for predicting their behavior switch right at that moment when they come out. Before, they come out, the soil temperature is important; after they come out, air temperature and the weather is important. And the farmers are hoping that you'll have a cold snap and all the grasshoppers will get sick and die, and the crops will be safe.

So, the trouble was that in our data, we didn't have an observed date for that hatching date but we knew sort of qualitatively what it had to look like and so we hypothesized a set of features that would let us predict a predicted hatching date and use that as part of the model. So, this is a case where the features we designed were based on a qualitative causal model of the domain. And yet in the final system, they are certain to be invisible; you just have some set of features about the weather and another set of features about the weather and some weights and some thresholds and it's learned there.

Satoru Miyano:

I am interested in the role of humans. You mentioned data mining and learning systems. A naive interpretation is that they are not just users but somebody who gives raw data. But the more important thing should be the interpreted data. There are two questions, one is how much of the information from the data bank or machine learning system can be really adapted? And I am also interested in the teamwork between the machine learning system and human team. So, for instance, one of my friends is working on detecting symptoms of air terrorism from network information and he emphasizes the teamwork between the machine learning system and human team is very important. But it seems to me he has not answered and he left those problems for the future.

Thomas Dietterich:

I agree completely that right now these are all man machine systems and particularly, I think this is true in a lot of data mining applications because of what you're looking for in data mining is usually this so-called actionable knowledge. Well, the machine learning system doesn't know what that is. There is some model that the human has about the business, what are the actions the business has available to it, what is their business plan, the relative costs of the different things. And so, interactively, working with the data mining system, humans try to find rules, knowledge, regularities that they could exploit in their business plan. What I would like to see is that we represent that business plan, that we understand what the business of the telephone company or the e-commerce company so that the machine learning system doesn't output silly obvious things. The machine learning system rediscovers things like there are men and women in the world and because they come out of the statistics and they are reliable knowledge, there is no question about that. They are statistically reliable but it is something that is completely obvious. And someone was talking earlier today about having a system for filtering those out of the discoveries. And if the learning system had a model of the business, it would know what was relevant and what wasn't. That's not a complete solution but I think it points to the reality.

SESSION2: Discovery

"SD, ML, KD, KA"

Hiroshi Motoda

ISIR, Osaka University

The Future of AI Workshop

December 14 – 15, 2002

AI: Learn and emulate Human Intelligence and apply it to problem solving

- Feel of understanding
- Proper level of granularity of reasoning steps and concepts used
- This applies to SD, ML, KD and KA.
- Increasing complexity and volume of data send a challenge.

Scientific Discovery

- Beyond system identification (LSI, NNN,GMDH,…) and constructive induction by heuristic local search (BACON, FAHRENHEIT,…)
- Computer assisted discovery of first principle
  - Notion of scales, theory of regime and ensemble
  - Discovery of simultaneous equations for both active and passive measurement data
- Challenge
  - Discovery of dynamic differential law equations

- Impact
  - Gives deeper insight into complicated phenomena, where domain theory is not well established and much depends on data   e.g. non-physics domains such as sociology, psychology, physiology, etc.

Machine Learning

- Identifying right primitive descriptors
  - Feature construction (discovery)
  - Current approach is heuristic/opportunistic.

- Coping with large volume of data
  - Feature/instance selection (reducing the data)
  - Incremental boosting
  - Active learning

- Maintaining comprehensibility
  - Much of the efforts on increasing predictive accuracy
    - Ensemble/committee learning
    - Meta learning   ⋄ Use of unlabled data

- Comprehensibility of a single hypothesis lost
- Efficiently using learned knowledge
  - Utility problem solved?
    - Our experience of geometry problem solving
  - Meta learning helpful? → Same problem of KA

Data Mining

- Mining from structural data: elements of data related to each other, e.g. chemical compounds, web log, patient records
  - Graph mining: Finding interesting/important patterns
  - Challenge: Subgraph isomorphism known to be NP complete
    - Levelwise complete search using graph invariants for frequency measure
    - Heuristic approach for non-monotonic measure
  - Impact: Enables to explore new application areas such as drug design, evidence based medicine, etc.

- Enhancing communicability of mined knowledge
  - Mining knowledge of the form domain experts use is quite important.
  - Active mining: let experts be involved in all DM phases (data gathering, mining, evaluation – intensive interaction with domain experts is a key to success.)
- Counting everything really important?
  - Sometime more appealing to show one counter example with what is well known and understood in the domain

Knowledge Acquisition

- Acquiring knowledge without pre-modeling
  - KA bottleneck embedded in the knowledge modeling
  - Rather build a model after KA, e.g. recent RDR research
- Integrating KA from both human experts and data
  - Simultaneous/sequential KA from different sources
  - Different communities for ML and KA

Conclusion

Scientific Discovery

- Discovering First Principle Laws

Machine Learning

- Identifying right primitive descriptors
- Maintaining comprehensibility

Data Mining

- Mining from structural data
- Enhancing communicability

Knowledge Acquisition

- Integrating KA from both human experts and data

Session 2 Panel 4 Discussions

Edward Feigenbaum:
Would you say some words about "Different communities for Machine Learning and Knowledge Acquisition" in your presentation?

Hiroshi Motoda:
By knowledge acquisition here I mean acquiring the knowledge from human experts. There is a KA community headed by Brian Gaines in Calgary who used to have yearly workshop more than 15 years (this workshop has recently turned into knowledge capture conference). There is also the European community which has been organizing a similar knowledge acquisition workshop for many years. They recently discuss ontology stuff, but has long been discussing how to conceptualize the human problem solving activities and build a knowledge-based system in such a way that the knowledge or the structure can be shared for other problems. However, when you go to the machine learning conference there is no such presentation as far as I see. There should be some communication between these different types of people.

Thomas Dietterich:
And I think sociologically it's because the machine learning people are trying to automate away the knowledge acquisition. They want fully automatic systems. But they are really doing the knowledge acquisition. They just don't admit it. That's what featuring engineering is. But the knowledge that's captured isn't captured. It's just used in the engineering of the features. It's not represented in a way that they could use later.

Stuart Russell:
I think one reason for that is that there hasn't been work on developing representations capable of capturing—and experts believe that a certain feature is relevant, but we don't know in which way it's relevant.

Manuela Veloso:
I thought discovery would involve also this kind of preparing your mind to get surprised. So I'm wondering whether we will ever be ready to have an AI system that will be capable of being surprised. Or what does that mean? Learning can be recognition, or pruning of relevance, finding relevance, finding causality. I have a whole book on serendipity and discovery associated with being surprised by preparing your mind to be. So what does that mean in terms of our systems, or in terms of our theories?

Yukio Osawa:
In Japan we had a very big snow last week. My student's client won 50 billion yen for that snow because we were just using WEKA. WEKA is a free software. So we used these a priori methods for the first learning of association rules. This is just an ordinary learning system. But the point was in the post-processing of data mining. The maker of the car tires, they accepted our knowledge, our discovery, that we predicted that in the middle of December or in the beginning of December the first big snow will come down in Japan. So that's our prediction. They accepted this. It was a surprise but they accepted this because we previously predicted that October will be the first snowfall in Hokkaido. That came to be true. And that is why they accepted the new predic-

tion. So the point is that the human to human trust is important. We discovered the new knowledge from using WEKA, but it's a free tool. We have established the trustworthiness between us through human to human interaction.

Takashi Washio:
But I think that the trust between human and the discovery is not the same.

Yukio Osawa:
That is not the same at all. I am just saying that these are two separate things, but trust is sometimes necessary.

Manuela Veloso:
I was less ambitious in my question. I was just asking to be surprised by a physical phenomenon, to be surprised by something. If I'm given many images of satellites, would I be able to find a new satellite that nobody knows about? Would I ever have a system that hypothesizes the complement of what it's given, or is the complement something we can also model, and therefore it's also a feature?

Michael Witbrock:
It's inconsistent with your current model but it's not probably false.

Manuela Veloso:
So I guess we just label it as a non-hypothesis.

Eric Horvitz:
I believe that an interesting challenge area centers on the construction of models of surprise. An important part of work on this challenge is defining "surprise." There are statistical measures of surprise, and as well as other notions of unexpected outcomes. There are opportunities to develop methods for reasoning about surprise, including approaches that abstract sets of surprising outcomes into an ontology of classes of surprise in a variety of arenas.

Manuela Veloso:
I have a book about 25 or 30 of these discoveries. There is something that the data might be present for me and for you, and you yet have your mind prepared to be surprised and I don't. And I'm wondering why, what is that little bit of something that prepares your mind for the surprise?

Thomas Dietterich:
I just want to speak on behalf of Herb Simon who is not here unfortunately. He says that the person that is prepared isn't really surprised by this. So, he would be very skeptical of these claims of serendipity in some sense.

Kumiyo Nakakoji:
I have a rather philosophical question. If we have super intelligent machines, then we cannot really tell if they are super intelligent because it is beyond our intelligence. And this nice discovery thing may suffer the same thing if the machine discovers really super knowledge beyond our knowledge. Then how can you really or how can we really see it as knowledge? So, what is the

definition of knowledge? And maybe there is useful knowledge for the machine but not for the human beings.

Einoshin Suzuki:
In the data mining community, there is a group of people doing exception discovery. So, one branch discovers exception to data, another branch discovers exception to knowledge but the third branch including discovers patterns which imply the exceptions like some strong regularity and exception rules together.

Takashi Washio:
The exception is it has to be only one example of discovery. The discovery community is not trying to automate the discovery. We essentially assume the existence of the human in the process of discovery. The discovery system just supports the humans.

Kumiyo Nakakoji:
If that is the case, then the definition of knowledge is understandable for humans and not for machines.

Manuela Veloso:
There is a lot of depth in your question so are we all agreeing that an AI system will never come out with something that we will then struggle to explain or some discovery that someone did and not the rest of humanity did? And then this person has the trouble of explaining what it is so that we all now know what the discovery is.

Eric Horvitz:
Automated perception and reasoning systems have already yielded such discoveries. Let's take an example from histopathology. Histopathologists examine cytological features in tissue sections to make diagnoses. Several years ago an image-processing research group examined a set of cases of lymph-node sections and employed vision algorithms to go over them and evaluate, among other features, the statistics of patterns of edges of chromatin that were not explicitly considered by pathologists. As I recall, a number of these features that were easy for the vision system to see but unnatural for people were valuable in boosting prognostic forecasting. Such findings lead to new questions and point the way to new knowledge. For the case discovering new visual patterns by automated image processing, we might ask the question: What are these new chromatin patterns being discovered in the nuclear compartments of these cells by the machines that people never noticed? And that could lead to a new line of investigation. So, overall, there is great promise for intelligent machines and analyses serving one day as partners in scientific discovery and confirmation.

SESSION3: HCI

"Capturing, Processing, and Exchanging Interactions Situated in the Real-World"

Yasuyuki Sumi

ATR

The Future of AI Workshop

December 14 – 15, 2002

Past Works: Support of Interactions

Externalization of contexts in a group

Visualization of the group context

Casual

Agents participating in human conversations

Turning the individual/group contexts into a story representation

Personal agents

Collecting real-world contexts

Real-world
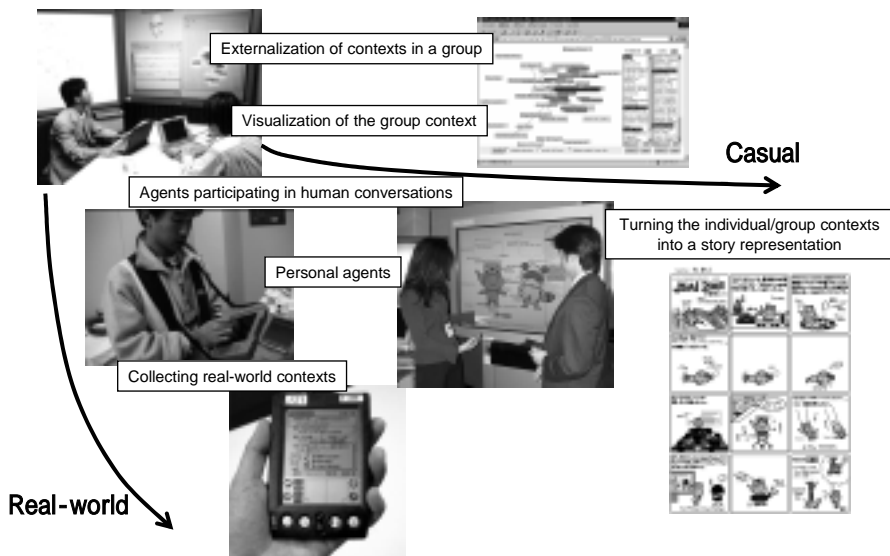
Interaction Corpus as a HCI dictionary

Multi modality
(including non-verbal)

Tacit knowledge, skill   Everyday interactinos

Spoken language corpus

Sign language corpus

Ontologies (for expert systems development)

Bilingual corpus (for machine translation)

CYC, WordNet, EDR, GDA

Mono modality
(Verval-centered)

Special knowledge
(Textbook-like knowledge)

Common knowledge
(Everyday knowledge)
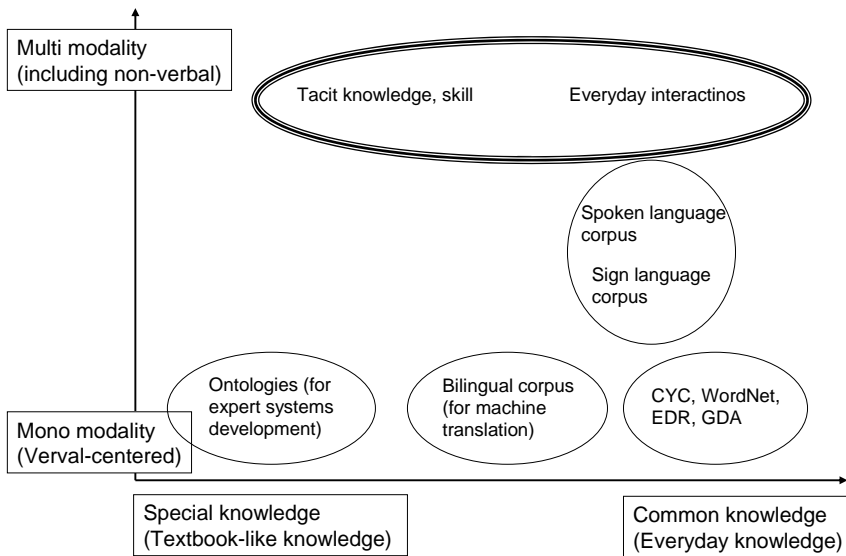
Capturing Interactions

- Ubiquitous + Wearable sensors
- Target detection by LED tags
- Co-creative partners
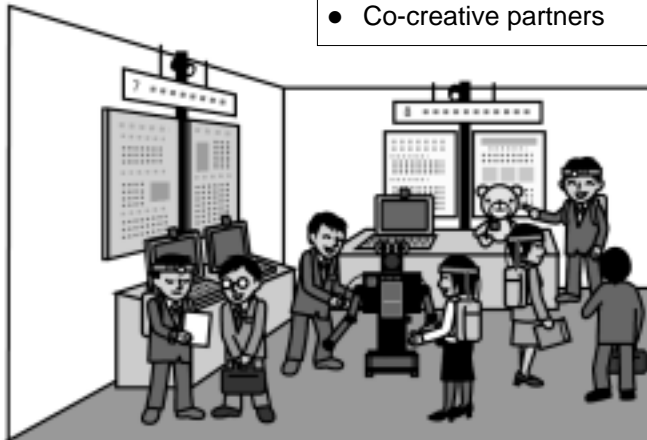
Capturing Interactions: First trial

Ubiquitous sensors (camera, microphone, LED ID sensor)

LED ID tags embedded in the environment

Camera, and LED ID sensor & tag

Adjacent microphone

Note PC inside the bag

Automated video summarization

Extracted scene
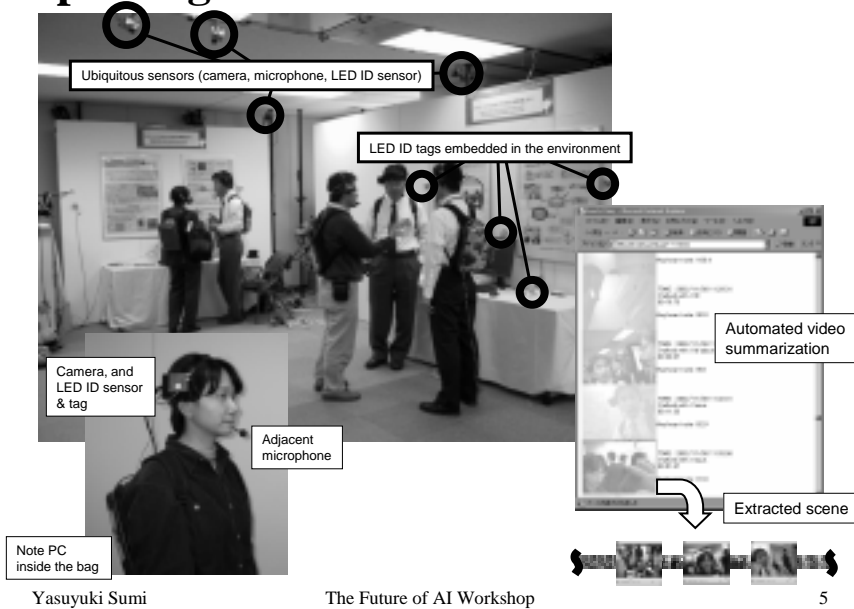
Towards the Experience Web

Humanoid robot

Ambient interface

Real world

Experience facilitator

Interface agent

Interaction corpus

Interaction dictionary

Community corpus
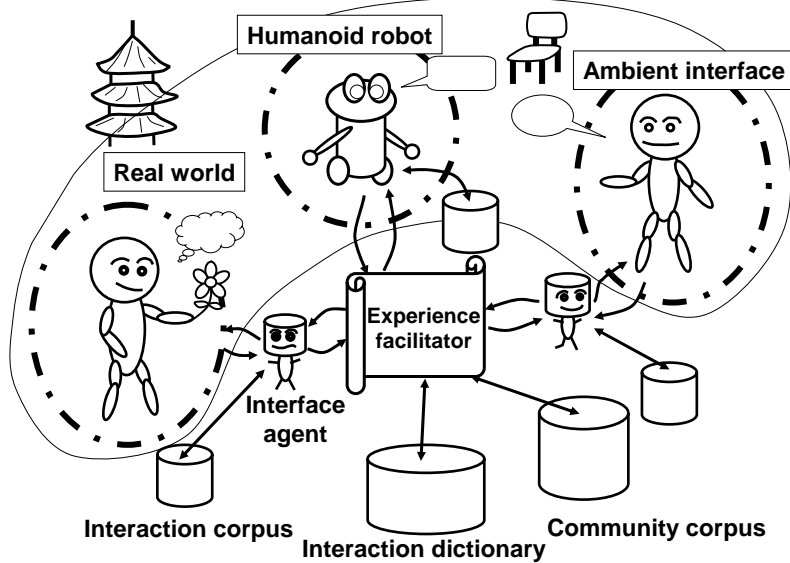
Session 3 Panel 1　Discussion

Yukio Osawa:
How do you distinguish between co-creative partners and not co-creative partners?

Yasuyuki Sumi:
This is very difficult, but at least my conclusion is that we'd like to use active partners, active artifacts, for making experiments, not only positive capture by cameras and microphones. In the future we will have to travel between many distributed artifacts. So in that case, we will not have the usual, conventional interface.

Yukio Osawa:
I intended to ask you about the quality of the partners. So, if you see someone else, then you may feel some surprise because his or her interest is quite new to you. But in some situations you can accept the surprise, but in other situations you cannot accept it. So the timing, and the situation, and context may vary the co-creativeness of the partner.

Yasuyuki Sumi:
Currently, I didn't think about the partners for providing some new findings to the individual users. So, my intention was only on the capturing side. So using only positive capturing devices it is very difficult to experiment interactions. It is very difficult to define the primitives of the interactions. But by introducing such co-creative partners there are some intentions, machine-readable intentions within such co-creative partners. So it became easier to segment our interactions. So this is our intention.

Eric Horvitz:
It wasn't clear to me exactly what the end use of the analyzed corpus will be. You mentioned some about what a user would see, derived from the captured materials. How will they be used? How in particular are you going to use the captured data? What are the applications?

Yasuyuki Sumi:
The big goal for us is to create a symbiotic situation between humans and artifacts.

Eric Horvitz:
You've got a real-time application, and I imagine you can probably develop many applications in that realm. Then you have applications that are designed for the offline review of the material.

Yasuyuki Sumi:
Currently, as a trial for starting the interaction corpus, it is offline. But by analyzing the pattern of the interaction, social interactions between users, between humans, and humans to artifacts we will have some kind of dictionary for interaction, including not only text-type or keyboard and mouse-type interaction, but also quality expression, with the goal of enhancing future real-time systems.

SESSION3: HCI

"Interaction Design for Software Aesthetics"

Kumiyo Nakakoji

University of Tokyo

The Future of AI Workshop

December 14 – 15, 2002

---

1. Historical Background

Systems that support people in designing

1: *Critiquing systems*
   a user talked about the system's suggestion

2: *Systems for creative thinking*
   a user became aware of a new aspect, but the design flow stopped

3: *Systems that do not obstruct people's thinking processes*

2. Interaction Design

industrial design for software systems [Weed 96]

selection of behavior, function, and information and their presentation to users [Cooper 99]

not only
- interface design
- graphic design
- information visualization

Interaction designers determine the inside of the product by describing the outside of it

3. A Story

I need a help in drafting a proposal!

How about sketching for text editing?

Possible AI solutions:
1. spell checker
2. grammar checker
3. thesaurus
4. semantic analysis
5. text generator
6. automated organizer

4. Toward Support for Interaction Design

Process Architecture

Interactive Systems
ART#001, 002, …

end user

interaction designer

Interaction Method
spatial positioning for
linear information design

EV ES
EE

programmer

Design Principle
ART

Software
Component
Library
Jun-NSN

Kumiyo Nakakoji   The Future of AI Workshop   5

---
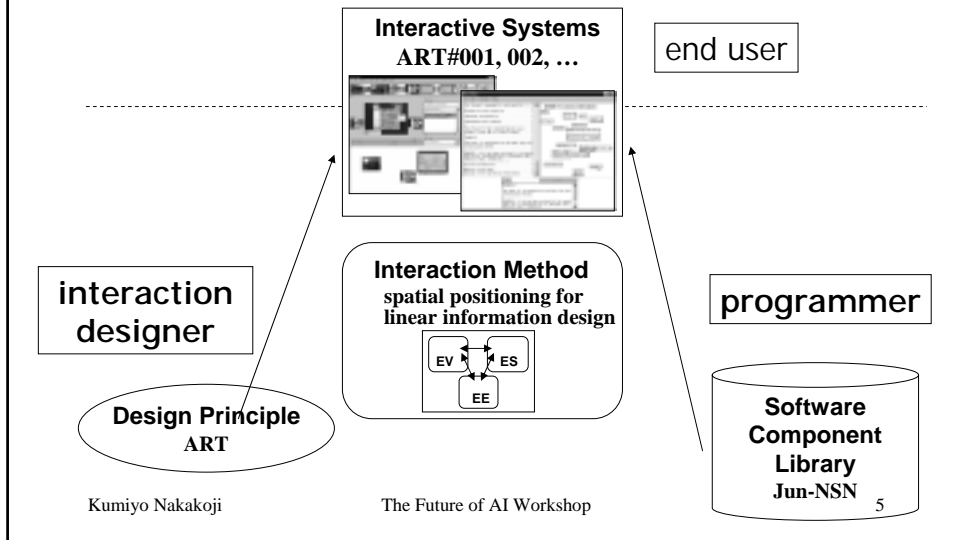
5. Conclusion

Software Aesthetics

*"Aesthetical point of view basically is a logical question,
not primarily a question of psychology, ethnography, sociology,
etc. It is a basic axiom here that it is through the force of its
inner logic, its consistent appearance, that a thing receives
depth in its expression and thus its strength to act as a
placeholder for meaning."* [L. Hallnaes, J. Redstroem, 2002]

Interaction design

- rethinking software development
- transcending the tradition

    not providing a software solution to an application

    but resolving a user's problem into a new task

Kumiyo Nakakoji   The Future of AI Workshop   6

Session 3 Panel 2   Discussions

Daniel Bobrow:
The issue is not what the system does or what the human does but what the team can do together. And apart from what you're talking about here, how do you tighten the team boundaries so that they are doing what they need to do together? So saying it's not the back-end does a different kind of injustice, because what you've really got to do is get the entire function going. And I think that interaction design is about the process of the team working together to achieve a goal. And you only implicitly talked about the goal before.

Kumiyo Nakakoji:
Yes. In fact, we have been looking at how this kind of interaction design can actually be performed. And I have observed the collaboration between interaction designer and programmer for the last three years. And there are really interesting things going on because the interaction designer often comes up with a nice interaction design idea, which has to be implemented, and be executable on the computer system. So while designers come up with desirable factors for the systems, programmers, or technologists, or engineers can only come up with what they can do. It's like making a lot of compromises between those two groups of people. And also, there sometimes are things an interaction designer has never thought of before, which can be achieved with the current CPU and the hardware technology, a programmer comes up with a cool kind of visualization that changes the interactive design. So it's a really nice synergy going on, and I hope that more of this kind of collaboration shows up.

Eric Horvitz:
You want designs that don't get in the way of the creative process, or if they do provide creative input should not be disruptive to the creative flow. How does one do that?

Kumiyo Nakakoji:
We design it under certain principles. Like in our particular project, this designer really influences the power of visualization. And he calls it "representation of talkback." So the visual appearance really matters. And he has put a lot of emphases on how the visual input and visual feedback is carefully designed while not ruining the perception time. If the CPU takes much longer than the interaction designer expected, then he just abandons that option and comes up with another option. But what I would like to stress here is we don't have this kind of design scheme or principle in computer science or on a computer display yet. And, for instance, if we look at the automobile design, there is a Porsche, and a Benz, and a recreation vehicle. Each has a very different value with a different lifestyle and a different emphasis on it. But so far we only have one type of software system that can be used by everybody for everything, and I think that we have reached the point that it should not be like that anymore. Users have to be able to choose the options that fit their style. So in using the systems that we have been developing some people really love them, but some people hate them. And I think that's fine because design cannot solve all the problems. Systems should be designed for a particular kind of person with a particular style of making process.

Hiroki Arimura:

I agree with you in saying that the interaction here is quite different from what will be written in the proceedings afterwards. So the question is, how can you think about the implication? What can we do for that? Is there any means for us to communicate the discussion here to the people who could not come?

Kumiyo Nakakoji:

I think it's a really big research challenge, and we should carefully think of what we would lose by having a different time scale in interaction. And there are many other factors, like awareness. There is short-term memory and long-term memory, and during the discussion we mostly use the short-term memory, but in the written material we do not. So how can we complement that with the computation media? For instance, instead of just reading, the computers could demonstrate the text by different colors as time elapses, or different animation can be used in the written material. So such a kind of visualization is going on.

Hiroki Arimura:

If I formulate my question more formally, the question might be put in this way. What can be involved in interaction? Can we have some means to represent what is involved in interaction? Probably there are several kinds of things involved in interaction, just for communicating purposes. That kind of thing may not be important later, because maybe that is a controversial issue. But maybe just for communication, I wonder if there is something else very essential that can only be communicated by this kind of interaction.

Kumiyo Nakakoji:

I think that we have to look very carefully into sociological issues. I like the idea of a socially shared, communicative environment when you are communicating with somebody else you share this environment. It's a communicative environment. And with the written material we need to use such a kind of communicate environment. I don't think we have a good model or representation for describing the interaction or the interactive environment yet. That might be one field that we should work on.

SESSION3: HCI

"Intercultural Collaboration Experiments 2002"

Bringing AI and CSCW Together

Toru Ishida
Kyoto University

The Future of AI Workshop
December 14 – 15, 2002

---

1. Background

- Asian Internet population will surpass US and Europe
  - Asian Internet Population in 2000
    - 78 million
  - Estimated Asian Internet Population in 2003
    - 183.3 million
    - 162.8 million in the United States
    - 162.2 million in Western European Countries
- A language barrier exists in Asian collaboration
  - Global collaboration is ongoing in English speaking countries
  - European Union encourages Europe wide projects
  - Asian engineers/researchers want to make documents in their mother languages
- Real issues for AI and CSCW
  - Multilingual and Intercultural Collaboration

Toru Ishida   The Future of AI Workshop   2

2. Experiments

Intercultural Collaboration Experiment 2002

- Open source software development in Asian countries
  - Shanghai JiaoTong University (China), Seoul University, Handon University (Korea), Malaya University (Malaysia), Kyoto University
- Development in <u>your MOTHER language</u>!
  - Grounding to common languages, i.e. programming languages
  - Team members never see in person, and complete software with multilingual Communication tools: TransWEB and TransBBS.
  - Each country develops a multilingual tool
    - Japan: TransGroupware
    - Malaysia: TransSMS
    - China: TransSearch
    - Korea: TransChat
  - April-June: <u>Software Design</u>
  - October-December: <u>Software Integration</u>

3. Findings

- Participants had difficulty understanding the meaning of translated messages/documents.
- Adaptation emerged…
  - when writing
    - Translation tests before posting again and again
  - when reading
    - Browse multiple languages on TransBBS
- Collaboration emerged…
  - for increasing translation quality
    - Communication to confirm message meaning
  - for overcoming translation errors
    - Technical discussions regardless of poor translation quality

4. Bringing AI and CSCW Together

- Realize a translation agent
  - as a social entity in a multilingual project
  - by combining AI and CSCW technologies
- Support participants to adapt to machine translators
  - for both writers and readers
- Support participants to overcome translation errors to develop logical discussion
  - Collaborative translation refinement by analyzing *repair organizations*
- Design protocols for participants to achieve communication even if they use noisy media
  - *Collaborative ontology* by adding tags on posting messages

5. Conclusion

- Bringing AI and CSCW together
  - To solve real issues in a multilingual project
    - Machine translation is one of the achievement of AI
    - Intercultural collaboration is one of the major issues of CSCW
- Future Schedule
  - ICE2003: Multilingual Development of Translation Agent
  - To develop multilingual collaboration tools in five years so that Asian people can freely create joint projects.

Session 3 Panel 3   Discussions

Daniel Bobrow:
It looked like this gave a drive to create new, very interesting ontologies. It also, perhaps, gives you some tools—or could give you some tools for doing these ontologies. Have you thought about, for example, the examples of use, what else looked like this before, et cetera? Are you doing anything with this corpus in order to help build the ontology?

Toru Ishida:
We don't think that a corpus is helpful in this case. When we started the project, there was no data. But the project always created a lot of terminologies, and so we tried to create tools for people to collaborate with ontologies. I don't think we can make a large corpus for short term projects.

Katashi Nagao:
I'm interested in the machine translation part in your project. Machine translation mainly fails because of ambiguities included in the sentence, for example, structure, syntactic ambiguities, and semantic ambiguities. In that case, if the user adds a marker for more sense annotation, then it improves machine translation result very much. So which is more difficult for an ordinary user, to mark a text or to rewrite that text?

Toru Ishida:
The students started to use their own markers for their language. I suggested to use Hasida-san's GDA. They tried to make an annotation to specifically define semantics.

Katashi Nagao:
There are tools to make semantic mark up easier.

Toru Ishida:
Yes, we need such tools. This event has just ended last week, and we are starting to analyze the data. We have thousands of messages, and probably several thousand of repairs. We will try to analyze that data.

Manuela Veloso:
What's the use of the English? I mean, you just look at the English?

Toru Ishida:
English is very important. For example, I write in Japanese and I translate it into several languages, then I only understand English. I check English translation, and if it seems okay, I send it on.

Manuela Veloso:
So if it is not, then you say it in a different way until the translation in English captures what you want to say? So you are trusting that the English to the other language does it right, right? So when you say "eight-times interaction", it's someone refining their text eight times. But who refines the other side?

Toru Ishida:
We have a translator from Japanese to Chinese, Japanese to Korean, Japanese to English. But we cannot know the quality of translation between Japanese and Korean, Japanese and Chinese.

Manuela Veloso:
So you don't know if this is what you're getting.

Toru Ishida:
No. English is all the information source.

Michael Witbrock:
So you believe there's reason to hope that their translations are right, and the other ones are not, and so forth. Is there any justification for that hope? Is there anything similar about the translation systems?

Toru Ishida:
Usually Japanese to Korean is much better than Japanese to English, and Japanese to Chinese is a little bit worse than Japanese to English. This is our experience, but we really don't know if the message is really being translated well.

Daniel Bobrow:
But do things that go wrong in one translation also go wrong in the other one? If you have a mistake in English, do you get a similar mistake going to Chinese or other languages?

Toru Ishida:
We don't know that.

Stuart Russell:
There are several ways that such mistakes could occur. They could occur in the understanding of the original sentence, and then also in the generation phase.

Naomi Miyake:
I think the whole key in this project is that the people in this project are sharing a goal, and they know what they are talking about. And for this you can take that into some account to start teasing your data. I don't think you can get the whole thing out.

Manuela Veloso:
But then I'm curious, why does this person have to refine each query eight times?

Naomi Miyake:
My guess is that for the first time, when they are trying to match the original language or the mother tongue to English translations they may start worrying about the small details about the grammar and so on, but if this has lasted long enough they would stop worrying about those smaller things and start seeing what kind of reaction you get. And this is not for the communication's sake, but they are developing some common thing. So if you say "hurry up", and if the thing comes quicker than you expected you would know that what you wanted to say is commu-

81

nicated through. So it's all in relation to what you get from that transaction that you started to make happen and what kind of reaction you'd get in relation to the purpose that you are working in that situation. Those things would all interact with the language translation quality, and so on. In answering Nagao-san's question indirectly whether the mark up will help or not, if you want to have that kind of mark up support to facilitate the communication, I think you would need a socially action-oriented mark up language.

Toru Ishida:
We really need to semantic support. And the trick is that the semantics are in the level of programming. Students can share C++ code or some other programming result. Otherwise, it's really difficult to discuss and to decide everything using machine translation.

<u>SESSION3:</u> HCI

"Human-Computer Interaction"

Eric Horvitz

Microsoft Research

The Future of AI Workshop

December 14 – 15, 2002

Frontiers of AI in HCI: Human-Computer Collaboration under Uncertainty

➢ **Meshing learning & reasoning under uncertainty with HCI design**

   ➢ Fluid mixed-initiative collaboration

   ➢ Multi-channel, multi-timescale learning and reasoning

   ➢ Human-cognition—aware systems

Probabilities

• **User query**
• **User activity**
• **Content at focus**
• **Data structures**
• **User location**
• **User profile**
• **Vision, acoustics**

• **Preferences**
   Cost & value of actions?
• **Intentions**
   Needs and goals?
• **Attention**
   Focus of attention?

*
Utility-directed actions
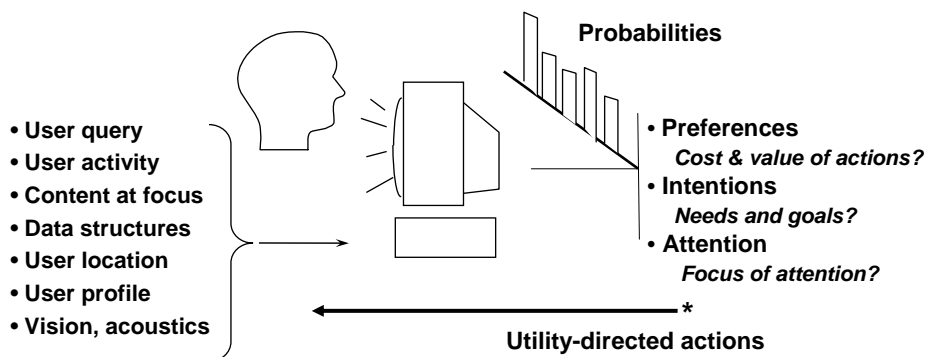
Toward Fluid Mixed-Initiative Interaction

◆ **Support interplay of contributions from human & machine to solve a problem at hand**

> ➤ **Architectures and languages for *grounding*—converging on *mutual understanding* of goals and context**
> ➤ **Reasoning about if, when, and how to contribute**
> ➤ **Problem recognition and decomposition**
> ➤ **Leveraging natural cues about intention and attention**

Automated grounding | **Mixed-initiative collaboration**
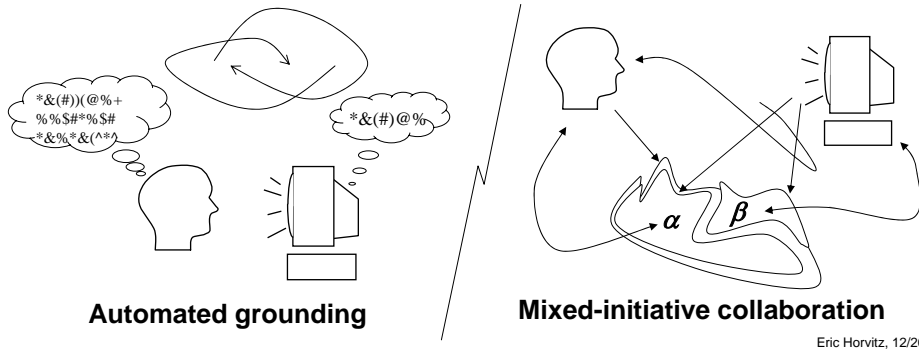
Multi-channel, Multi-timescale Fusion
for Identifying Context and Goals

◆ **Consider multiple sources of perceptual evidence**
◆ **Learn and reason at multiple scales of time**
◆ **Learning over extended periods of time**
◆ **Statistical fusion of multiple channels**

> ➤ **Build models to infer attention, intention, and goals**
> ➤ **Multi-channel data over extended periods for personalization**

e.g., Seer: **Layered HMMs for Office Awareness**
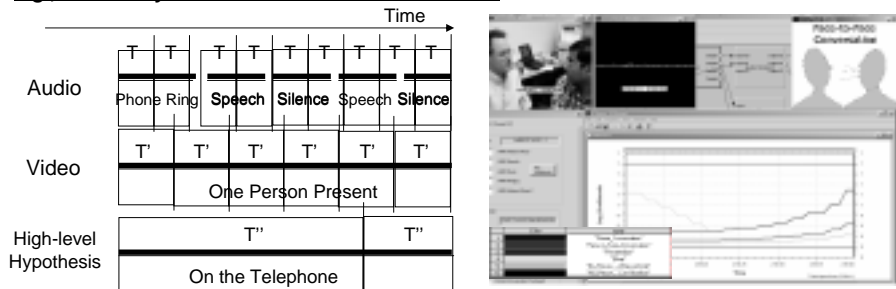
Human-Cognition—Aware HCI

- ◆ **Endow UI, services with knowledge about human cognition**
- ◆ **Apply results from cognitive psychology in design & real-time mediation, modulation, display**
  - ➢ **Divided attention and disruption**
  - ➢ **Memory and concept attainment**
  - ➢ **Visualization**
  - ➢ **Judgment & decision making**

Target Cognitive Tasks

* New Cog. Psych. Research

* HCI R&D

Memory   **Concept Attainment**   **Divided Attention**

Abilities & efficiencies
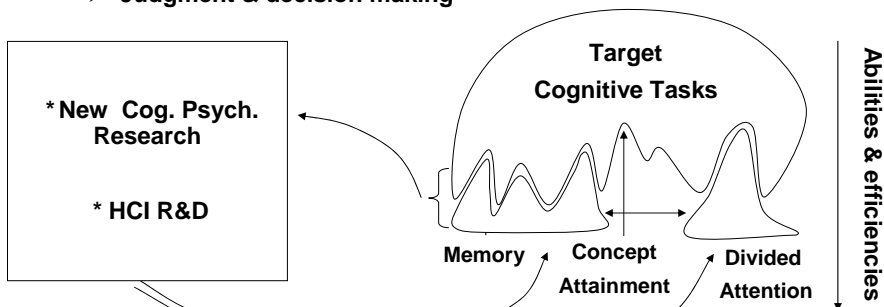
Eric Horvitz, 12/2002

Conclusions

Frontiers of AI in HCI:

New services and experiences by leveraging advances in learning & reasoning under uncertainty

- ◆ Deliberating about a user's preferences, intentions, attention, and context under uncertainty

- ◆ Architectures and methodologies in support of mixed-initiative interaction for fluid collaborative experiences

- ◆ Multi-channel, multi-timescale analysis for composing rich assessments of goals and situation

- ◆ Learning about users over extended periods of time

- ◆ Endowing systems with knowledge of cognitive attributes and constraints

Eric Horvitz, 12/2002

Session 3 Panel 4 Discussions

Edward Feigenbaum:
This understanding human cognition as a way of designing better interfaces—there's a famous example of this, although it's in the psycho-visual area rather than in the psycho-cognitive area. But color TV, the great break-through there was from psycho-visual people telling the engineers that people didn't use very many bits to see color. And almost all the information being transmitted was black and white. Suddenly the color video signal could fit in the bandwidth that was created by the FCC.

And that was not an engineering breakthrough. The trouble with HCI is it has been in the hands of engineers too long. The question I wanted to ask was for you, running a group at Microsoft Research for this area, do you have a two-year goal and a five-year goal? That is, an actual thing that you have your mind focused on, and you want that, and you'd love to see that in two years, and you'd better see that in five years, that kind of thing? At least ten years?

Eric Horvitz:
I continue to deliberate about a portfolio of research goals spanning different time horizons. I tell researchers on my team to pursue research that could have significant influence on the computing experience ten years or beyond. However, I also mention that it would be wonderful if such long-term research projects also happen to resonate with and have nearer-term influence on products and services. So, people on my team tend to be passionate about influencing products on a shorter-term basis while also pushing on longer-term dreams. Overall we pursue a mixture of challenges at multiple time horizons. We have several streams of work in progress, in communications, alerting, search, task management, and assistance. Broader themes include "augmented cognition" --reasoning about how we might best extend a user's abilities via machinery that understands various bottlenecks in human cognition uncovered by cognitive psychologists over the last 75 years. Beyond building reasoning machinery, such work may rely on the results of user studies performed in our user studies lab. As an example, we are working to build models and reasoning machinery that endow computer systems with the ability to deliberate about a user's attentional focus and workload, based on monitoring a user's activity and context. In this realm, we have been pursuing principles for trading off information awareness and disruption. We have built research models, one called the Notification Platform, employing Bayesian and decision-theoretic reasoning with long-term implications. We experiment with and demonstrate the proto-types as a vision to the company and research community about what may someday be available to all users. In the short-term, we field components, such as the Priorities email triage and mobile-messaging component-which shipped to the public in a product called the Outlook Mobile Manager. The prototypes have stimulated product teams to think about the right kind of basic abstractions, interfaces, and controls, and some of these will be appearing in products.

Manuela Veloso:
Are you telling me that you have algorithms that go from event logs to generalized models?

Eric Horvitz:
Yes. In fact, from the point of view of AI and machine learning more particularly, I believe we have introduced several innovations. For one, with the Coordinate system, we do not rely on

static Bayesian models that are built offline. In response to a query in real time, the system extracts equivalence class cases from a large event log, and builds a graphical model that is used for predictions about presence and availability. Typically the system builds several models within a period of a second or so, providing inferences to produce probability distributions over high-level events of interest. So yes, this actually works and it's not science fiction. And it's been a great AI testbed for doing this kind of dynamic model construction.

Manuela Veloso:
My event logs are x/y positions of where robots are moving, but I haven't succeeded in doing this. So there has to be some assumption in their mind for success. Where does the power come from?

Eric Horvitz:
Maybe we should take this offline for a more detailed discussion. I can give you a core dump on Coordinate and show you the details, including the simplifying assumptions we make. These include proximity, the way we're doing conditioning, and the kind of queries we allow—a set that is growing as the system is getting more sophisticated.

Manuela Veloso:
That's very beautiful. I mean, if you can't state the assumptions then you know exactly the scope of where this is feasible, which makes your answer being "this is not science fiction. This is real, using the assumptions that we made."

Eric Horvitz:
Now turning to my role as session chairman, we heard several talks each taking quite a different probe into the links between artificial intelligence and human/computer interaction. Are there any other comments on any of the talks?

Kumiyo Nakakoji:
I want to ask a question to you and probably to Sumi-san. I suspect that your model works very pretty if a person sticks to your model of living style. For example, once you put some calendar on your own note, then all of a sudden it breaks up, or something. My suspicion is that the system needs to be able to capture all the context of what's going on in the whole real word. And the model is based on the assumption to capture everything, and you put more and more computers everywhere so that the computer never fails to catch anything. But the other option can be a more forgiving system. It's okay that you forget one or two events that you kept a record of on your computer.

Eric Horvitz:
For people in the AI community it's very important to perform evaluation to gain an understanding of the power and generality of the principles employed, and to make sure that we're not fooling ourselves with specificity when we really should be solving a harder, more general problem. We do worry about these things. The types of queries that the Coordinate system handles is fairly robust in that it's reasoning under uncertainty, it's splaying out models that are probability distributions that capture failures and successes via notions of expectation, and it's exactly those kinds of inferences that decision tools downstream of Coordinate need to employ, such as components that compute the expected cost of delayed review in making decisions about messaging and communications. Those tools compute and reason about expected values. The expected cost, for

bothering Eric right now may be $4.85. The system computes that kind of thing. And the idea is to have a principled theory for how this should work under uncertainty. Your question to me is a good question but also suggests that you'd rather not face progress on a larger problem that admits incompleteness and uncertainty and that employs an explicit model for grappling effectively with such incompleteness. And I think if we're just clear about the assumptions that we're making we could make some progress and understand the borders of it and push out from there.

Stuart Russell:
It is actually a very broad problem in many systems that try to keep track of the state of the world. The assumption is always that for evidence variables in the past, the events that are supposed to be detected by the variables are detected. Your example that if someone put something in the calendar, they would put it in paper rather than putting it into your Microsoft calendar, does your probability model allow for the fact that someone has an item in their calendar but you don't know about it, or we have to assume that if there is no item on calendar there is no item?

Eric Horvitz:
That's a good example of handling potential incompleteness in observations. Coordinate allows for the use and testing different variants of the calendar analysis and how calendar information is folded in and we are experimenting with the different models. In one approach, information about a user's appointments is implicit and we consider patterns of presence, time of day, and so on, and sum over the calendar information. For a more explicit model, the system looks at appointment data and it assumes when nothing is on your calendar, it's going to be a default situation, learned for the cases where nothing is placed on the calendar explicitly. The default situation includes patterns of availability when people may have had appointments that are not encoded in their online calendars, which doesn't mean you're free because people do all sorts of things that aren't in their calendar.

We might find that particular forms of inference about a user's presence and availability fail in certain conditions. Such findings are valuable input for continued research on our handling of incompleteness or on the need to reduce the incompleteness via additional evidence gathering.

Stuart Russell:
People have this experience with inventory systems, that as long as the worker is diligent about bar coding any new items that are coming in, you keep track of your inventory book. In real life, in the sort of non-prototype experiment, then they're mostly just out buying five thousand new items.

Eric Horvitz:
Your other comment was, "maybe the outputs don't have to be so perfect." One of the applications we've been exploring is automated assignment of measures of urgency to email. I'd like to know if an e-mail sent to me is really an urgent e-mail. The Priorities system works to assign measures of the expected cost of delayed review to email and uses this measure, in conjunction with forecasts about my availability to determine when to transmit a message to me and/or when to tell the sender that I'm not available for some period of time. That is, if I'm not around the system may decide to bounce-back an out of office message that says, "Eric's not around now. I understand that this is very urgent, and I'll pass this to his cell phone. I expect him back within 90 minutes." It's telling somebody that they can expect me to be back reading email in about 90

minutes given my past behavior in similar situations, considering the time of day, and other variables, such as the type of meeting that I'm at, what kind of meeting it is, when it ended or will end, it's location, etc. But it's not sure.

Now, we can imagine less forgiving applications, where we would need to know an exact time, for some high-stakes coordination. In the applications we are targeting, it seems to be okay to make a best guess about an availability forecast based on previous experience, given the power of our inferences. The system actually states, "I'm guessing that…", or "I think he'll be back within 90 minutes." We try to mesh expectations in language with the accuracy of the system. My sense is that the system is a success if it performs as well or better than an insightful secretary in relaying inferences about the presence and availability of the person that he or she supports and understands who might need to know.

Edward Feigenbaum:
This is a question for all of the panelists, but feel free to answer it yourself if you want to. I'd like to find out the relation between advancing HCI, or using AI techniques to advance HCI, and knowledge-based approaches. You've talked a lot about calendaring, so obviously your HCI assistant here is using some rather sophisticated AI methods in the context of calendar. But I have a lot of contexts that I'm using when I'm working at a computer, and where am I going to get these knowledge models that are going to put everything in context, and inform the dialogue, and keep things on track, and so on? I'll be happy with this calendaring, but is calendaring one of three that you're doing? Or is calendaring one of a hundred, or...?

Eric Horvitz:
I don't view the Coordinate application as simply calendaring. We're looking at a very broad event stream, including what applications you're using. Coordinates also will tell us when you're likely going to shift from Word to other applications and when you will likely next read email.

Edward Feigenbaum:
And will the HCI agent on the Word side know a lot about my ability, the way I do text?

Eric Horvitz:
Let me step back to make a comment about the challenge of sensing and reasoning about the goals and needs of computer users. First, in pursuit of a more general science of this challenge, we've worked to better instrument applications and the operating system, as well as to introduce new perceptual sensors such as automated vision to recognize a user's pose and acoustical analysis to monitor ambient sound. We have worked to integrate new kinds of system monitoring, via creating or accessing hooks at a variety of levels of the operating system and applications. The legacy system and applications do not necessarily provide all the sensing we'd like. We developed an event sensing, logging, and abstraction system called Eve. Eve provides an event abstraction language. We can use the language to compile efficient policies for limiting or abstracting atomic events into higher-level events. We've been sharing key concepts represented in Eve with the product teams as a model of a rich eventing system that might one day be provided to developers, both internal and third party, for building new kinds of user modeling tools and services.

You come to realize, that, we acquire and need to analyze potentially large quantities of raw data -- just like any area of science where you collect streams of information. We acquire megabytes

of data from user sessions. The raw data includes every mouse move and click, and every word being typed. Eve also performs screen scraping. If we turn on screen scraping we log every bit of text in every window flashing up, and being scrolled.

Now, getting back to your comment—the challenge of developing models and methods for reasoning about users becomes like data mining in other areas of natural science. In the realm of human-computer interaction, we come to ask ourselves, "Okay, what's special here? What's the goal here? What are some applications here?" that guides our filtering, abstraction, and analysis of the massive Eve log, which also includes a user's calendar status and so on. So, in one application that we have pursued as part of our Notification Platform project, we seek to predict when a user might be back in their office again, or when the user might next read e-mail, or when the user might finish a conversation that is sensed as currently in progress.

But when we go to the operating system people and talk to them about prefetching and speeding up the OS to identify, say when a user is going to switch from Word to Outlook, it's a different application, employing different modeling methods and tools, but we might go to the same raw log, but filter it down and squeeze it down to a more compact, relevant log. So, we found that there's a gold mine of events and activity so large that you worry about keeping it around even on modern disk drives. So, logs of user activity and interaction are a rich arena for researchers to mine, especially in the context of attempting to field well-defined, compelling applications.

So I don't think there's anything different here from other areas of machine learning, modeling, and reasoning. In fact, it's more the same than different than other areas of science. Of course, there may be less regularity. That is, we might expect to have higher variances in people's work styles and so on than you would have with symptoms and diseases considered in artificial intelligence in medicine, or with molecules, in applications like DENDRAL for example. But in some ways, the high-level goals and challenges and principles for attacking the challenges are just the same.

Yasuyuki Sumi:
The intention of my presentation is to propose a big challenge for achieving the real symbiosis between human and intelligent systems situated in the real world. So in such situation we don't have the dictionary for understanding our human to human interaction, social interaction, micro-level interactions such as body gesture or non-verbal aspects of our interactions between human and human. And also my macro level, we don't have a dictionary or terminology—machine readable terminology for understanding or describing the micro-level of our social interactions, such as grouping few people in the open space in exhibition spaces or open social spaces, or joining or leaving some people in the group.

So then I'd like to propose the new challenge for building such dictionary, interaction dictionary, for understanding our human/human interaction. Then I hope intelligent systems will be able to participate in our human-to-human social interactions. So that's my intention.

Toru Ishida:
It's clear to me now that we need more knowledge support and more semantic support to create translation agents. And I think that machine translator is not enough for marketing of project. We need a translation agent which can understand not the details of conversation, but the role of each

message. For example, it should know if it is a question or an answer—if it should be done very quickly or not—or if it seems that some guy is getting angry, or something like that. To understand these, the meta-level knowledge of the message is also quite important. So we need two levels of knowledge support, a domain level and also a communication level in knowledge support.

Kumiyo Nakakoji:
I have a very different stance, and I'm kind of on the side of humans and regular users, so it's really exciting to see that machines can have such intelligence and supportive mechanism for me, and it's really exciting to see what the computers can do, but it doesn't necessary mean what a computer should do or will do for me. I think right now the danger here is, yes, it's nice to push the limits of the science forward, but at the same time humans have to be given a way to control over what technology to use. For instance, some people don't want to be monitored all the time. Then the big challenge comes into the HCI area, how can end users adjust, and adapt, and tune those behaviors? That's a really big challenge, and there is another whole set of AI technology necessary to support the users. So there is this ongoing kind of a fight—especially with Microsoft—but it's really a fight between users and the technology, and it shouldn't be that way.

Eric Horvitz:
Just one comment, here: It's very impressive to see how many user studies are going on everyday at Microsoft in user labs probing how people interact with a spectrum of products and features. It's interesting to drop by one of these labs and ask, "What are you guys studying today?", "We're studying Microsoft Money interface A versus B versus C for two weeks seeking user feedback about the different designs and experiences." So my comment to you, in reaction to what you have said, is: Look at these technologies as opportunities. Yes, they frame new design challenges for human oriented, human-centric design. Researchers and designers on my team and in teams outside of Microsoft Research on the product teams are not oblivious to the fact that people may want to control and custom-tailor configurations and behaviors. For some of the applications I have presented, yes, you want to have privacy controls, you want to have expressive controls for information alerting and flow. That's a whole different talk. Today, I really just briefly touched on the idea of what we can do with long-term learning from log of a user's patterns of activity. Now, perhaps it might seem that by not explicitly mentioning the other parts of the challenge of making the technologies usable and controllable that we're just going to attempt to transfer over the technologies to product teams and somehow persuade them to simply plug them in as is. In contrast, there's a lot of study on control and elegant design. Usability experts, and others, including people trained in subdisciplines like yours, are relied upon for their intuitions about if, how, and when these technologies are to be applied in real-world products.

Manuela Veloso:
So tell us an example of when things came from all these user studies and all these genuine efforts at Microsoft and went in to a product versus Microsoft thinking like this, "I have millions of customers. Guess what? Whether they like it or not, we'll shape the way they use computers," which is what I believe is the Microsoft approach, if I wouldn't know you personally. So I think Microsoft imposes on us whatever they care.

Eric Horvitz:

It's my understanding and experience that every Microsoft product available on the market has been studied quite deeply with users. It's important also to point out the distinction between Microsoft Research and Microsoft product teams. We work quite closely at times, but have different missions. I know well very diligent, focused psychologists on the product teams, and there's a lot going on with software evaluation. Human-centric design and evaluation is considered extremely important inside Microsoft.

Paul Cohen:

If the world's largest corporation, or second-largest corporation, has such difficulty designing useable interfaces, designing useable interfaces must be a very hard problem.

Eric Horvitz:

This may be getting a bit off-track from our themes today, but, in response to your comment, I must say that it has been interesting finding out from our colleagues on the product teams how hard it is to ship software that is used in very different ways by a large spectrum of people. We hear and sometimes see aspects of the process via our product-team colleagues, and learn what they go through in developing, testing, and shipping a product. It is indeed quite a hard challenge is to ship a great product that is used broadly, and, from the point of view of a researcher pursuing new core services with perception, learning, and reasoning, it's definitely a different world.

SESSION4: AI SYSTEMS

"The Challenge of Cognitive Systems"

Ron Brachman

Information Processing Technology Office

U.S. Defense Advanced Research Projects Agency
(DARPA)

The Future of AI Workshop

December 14 – 15, 2002

AI as a Systems Science

- Integration is the key
  - "Intelligence" comes from the collective abilities of the mind (doesn't reside in a single component or knowledge base)
    - Interplay between knowledge, reasoning, reaction/adaptation, environment
  - It's easy to ignore the complexities of integration in favor of tuning your favorite algorithm or data structure
    - Bugs and vulnerabilities at the interfaces
  - System architecture is underappreciated in AI
    - Need interaction with professional architects

AI Theory and Practice

- "In theory, theory and practice are the same; in practice, they're different" (Yogi Berra?)
- Practice matters as much to theory as the other way around
  - Our experience with CLASSIC: changes were needed to the logic to account for shortcomings unearthed by practice (e.g., rules, "test" functions)
  - Discoveries were made about theoretical complexity when practical challenges couldn't seem to be met
- Theory does matter to practice
  - Our systems are extremely complex and *ad hoc* methods are doomed to fail
  - AI is about generality and versatility

"It's no longer AI…"

- We need better applications/problems – those that intrinsically need generality to succeed
- We should also lay claim to "intelligent systems engineering" and take credit for successes that don't need generality but whose genesis is AI research
- Tempering of expectations
  - As always, we need to be careful of underestimation of the difficulty of our subject matter

DARPA's New Initiative in Cognitive Systems

- IPTO will create a new generation of cognitive computational and information systems with the capability to:
  - reason, using substantial amounts of appropriately represented knowledge
  - learn from their experience so that they perform better over time
  - explain themselves and be told what to do
  - be aware of their own capabilities and reflect on their own behavior
  - respond robustly to surprise

Systems that know what they're doing

Our First Focal Challenge:
An Enduring Personalized Cognitive Assistant

- Will have and use *knowledge* of the domain, task
- *Cognitive awareness*: will have experiences; perceptual input integrates with knowledge, model-based filtering
- Can *imagine* possible futures
- Can *decide* what to do and act in real time (prioritize)

Cognitive

- *Learns,* including by observing partner
- Can be *advised* and *guided,* and can *explain*

Personalized

- Must know how to *cooperate* (be a team player)
- Uses *multi-modal*, broad-spectrum interaction
- Should be available everywhere - *omnipresent*
- Must be *trustworthy*

Assistant

- Must *learn continuously*
- Must be able to *survive*, operate through problems

Enduring

Conclusion

- AI needs to be thought of as a systems endeavor
  - Integration is key
- Theory and practice are essential complements
- Applications that essentially need cognition and generality are critical

SESSION4: AI SYSTEMS

"AI Systems interacting with human"

Takashi Washio

I.S.I.R., Osaka University

The Future of AI Workshop

December 14 – 15, 2002

---

1. Intelligence does not imply communicability with humans.

Scaling up AI systems to supper intelligence and multi-tasking intelligence increases its functions, inputs and outputs.

Two ways to use AI systems,

   1.   Completely autonomous use without human interaction

      $\Longrightarrow$ Communicability with humans does not have to be considered in principle.

| Autonomous Missions in Deep Space |

   2.   Use in human society

      $\Longrightarrow$ It tends to decrease the comprehensibility for users.

Takashi Washio   The Future of AI Workshop   2

2. Increasing communicability with humans.

This is not only the problem of man-machine skin interface or media.

1. Background Knowledge

Socrates' dilemma

> How can we ever learn what we do not know? Either we already know what we are looking for, in which case we don't need to look, or we don't know what we're looking for, in which case we wouldn't recognize it if we found it. The only escape is to acknowledge that we already know what we need to know. (*Meno* 80e)

Large AI systems must share background knowledge with humans.

2. Reasoning and Interaction Process

Plato's claim

> All of us have had the experience of suddenly realizing the truth of something of which we had been unaware, and it does often feel as if we are not really discovering something entirely new but rather merely remembering something we already knew. The recollection may be the source of our true opinions about the most fundamental features of reality. (*Meno* 85d)

Large AI systems must conduct the recollection in itself for autonomy and support the recollection of humans through the interaction.

---

2. Increasing communicability with humans.

3. Objective oriented communication and Focusing the issues

The autonomy of large AI systems and its interaction with humans must have objectives in some sense. The objectives lead the reasoning and the interaction into some focused issues shared by both of the AI systems and humans.

Large AI systems must have objectives (motivation) shared by humans under the shared background knowledge and the shared recollection of the knowledge.

⇩

What needed to be solved:
• How do the large AI systems and humans share the knowledge?
• How do we measure the degree and the quality of the share?
• How do the large AI systems recollect required knowledge?
• How do the large AI systems settle and/or share the objectives and the focuses of issues?

3. Integration of different methods of problem solving.

1. The problem to share the background knowledge (e.g., axioms and assumptions) and to mutually support recollection of knowledge among the different methods though the interaction must be addressed.

2. The maintenance of the consistency among the shared/recollected knowledge, the objectives and the focuses of issues must be addressed.

These considerations suggest many new research topics.

2. Conclusion

Communicability Issue
1. Large AI systems must share background knowledge with humans.
2. Large AI systems must conduct the recollection in itself for autonomy and support the recollection of humans through the interaction.
3. Large AI systems must have objectives (motivation) shared by humans under the shared background knowledge and the shared recollection of the knowledge.

Integration Issue
1. The different methods must share the background knowledge and to mutually support recollection of knowledge.

2. The different methods must maintain the consistency among the shared/recollected knowledge, the objectives and the focuses.

Session4: AI Systems

Knowledge Founded BioMedicine
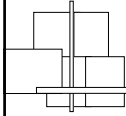
Koji Sasaki AdIn Research Inc.

Co Researcher : **Jun Nakaya** / Advanced BioMedicare Research Inc.

Acknowledgement for Simulation Knowledge: **Tetsuo Shimizu** / GMIN AMRC IMSUT

The Future of AI Workshop

December 14-15, 2002
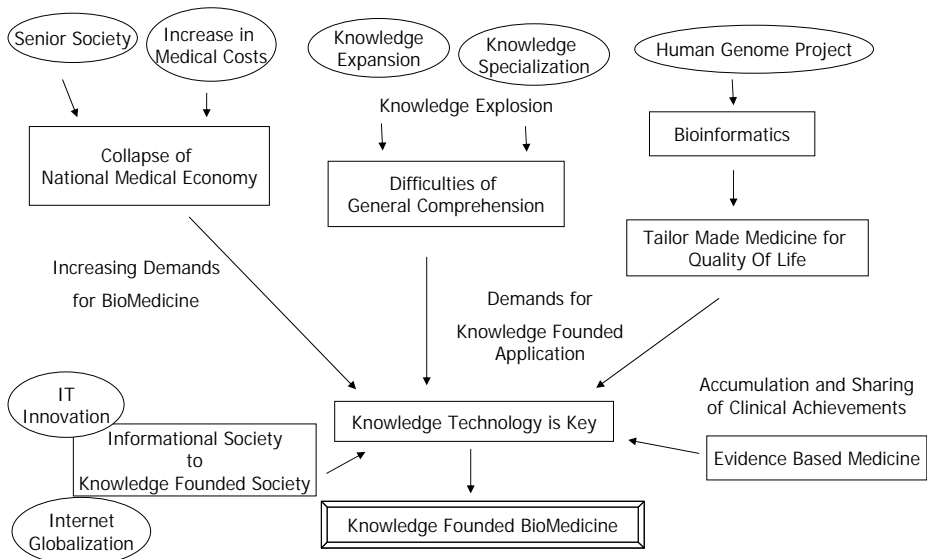
Age of Knowledge Founded BioMedicine

Definition of Knowledge Founded BioMedicine

Medicine & Biotechnology founded on Multi Disciplinary Knowledge

Examinational Data

Clinical Data

Statistical Data

Data

IN VIVO Knowledge
Macro Knowledge

Knowledge

Genome Data

Integration & Fusion of Multi Disciplinary Knowledge

IN VITRO Knowledge
Micro Knowledge

IN SILICO Knowledge
Theoretical Knowledge

Wisdom

Experimental Data

Simulation Data

Translational

Research

Research Issues
**Knowledge Description
Knowledge Acquisition
Knowledge Arrangement
Knowledge Utilization**

Knowledge Founded
Bio Industry

Knowledge Founded
BioMedicine

Drug Design

Tailor Made Medicine

Consideration of Knowledge Description Capability

IT
Oriented
Knowledge Tool

End Users
Oriented
Knowledge
Tool

**Natural to
BioMedical Thinking**

**Knowledge Plotting in
Multi Layer Multi Dimensional Space**

Description Capability

**Linking/Systematizing,
Integrating/Fusing
Knowledge**

Versatile Knowledge

Genome
and
Human

Biological Experiment
and
Clinical Treatment

In vitro & In vivo,
In silico & In vivo

etc
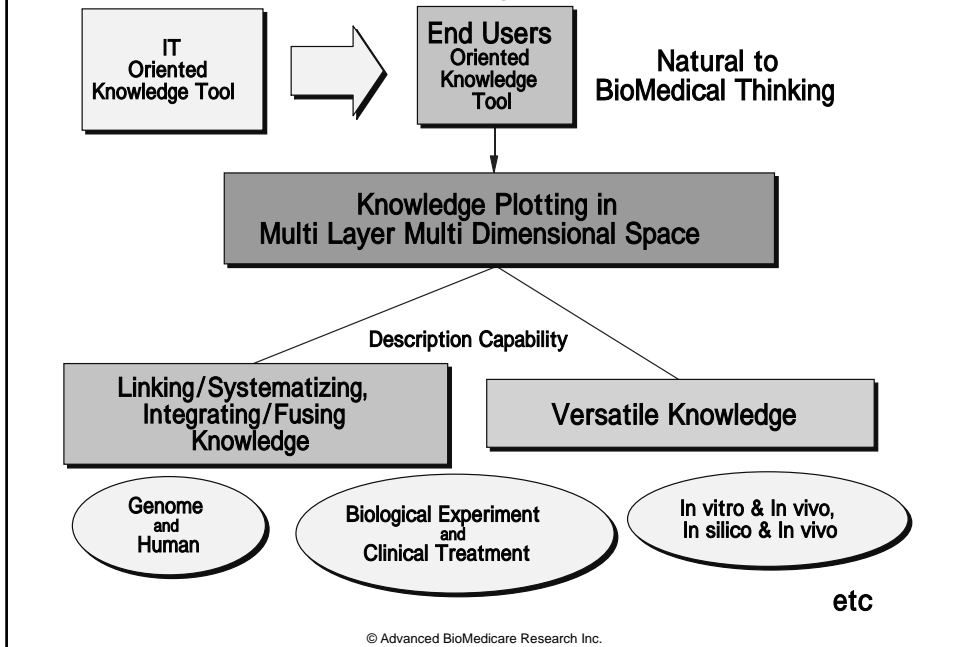
Multi Layer Multi Dimensional Space
An Example: Medical Space

Multi Layer Multi Dimensional Knowledge Space

Quantified Model

for "*Normal*" Range of Body Temperature

Medicine
includes

Diagnosis

Matching Degree

Pain

100% matching

Normal

Temp.

Area showing

Concept of *'Normal'*
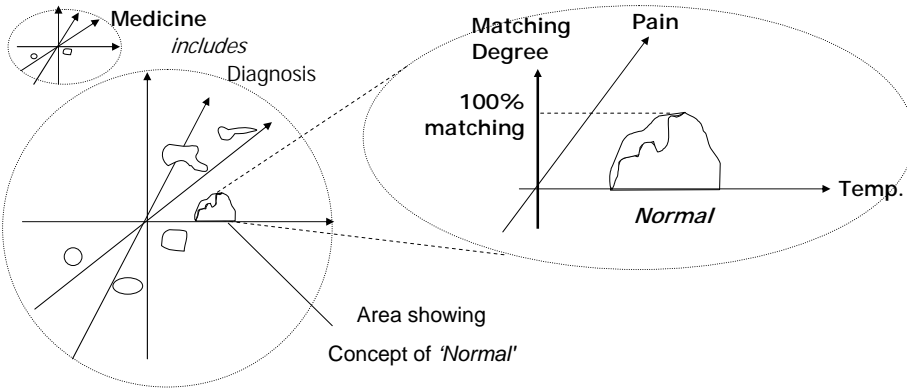
Knowledge Inspiring and Creation
Advantages of ML MD Space

Treatment

Treatment with Possible New Alternatives

Propose Candidate Area in Empty Space

Candidate Area

Inspiring

Candidate Area

Knowledge Space with Quantified Model

Creation

Extra/Interpolate with Fusion/Formation

New Treatment Formed
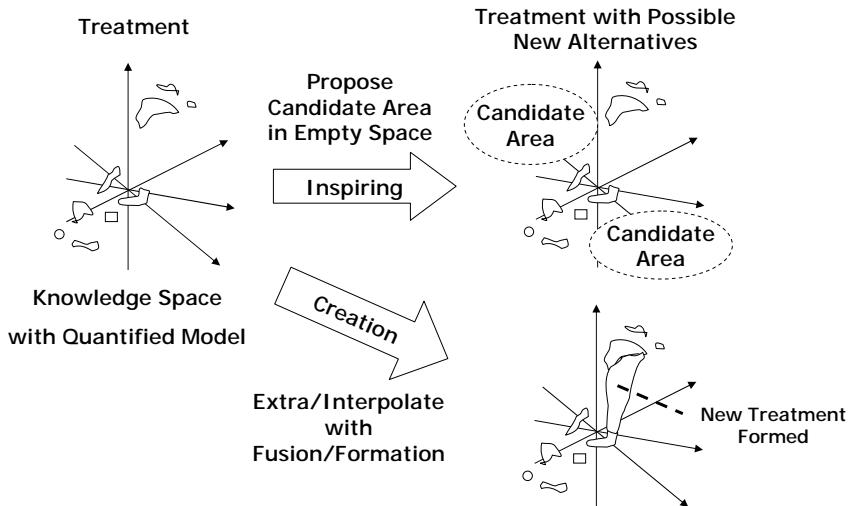
Conclusion

- Knowledge Founded BioMedicine is expected to overcome biomedical knowledge explosion in the age of fusion of IT, BT, and Medicine.
- Integration & fusion of multi disciplinary knowledge (in vitro, in vivo, in silico) is the key Issue.
- Knowledge Founded System must be users oriented and must be natural to biomedical thinking.
- Multi Layer Multi Dimensional Space can naturally express biomedical thinking.
- Description capability for versatile knowledge & linking/systematizing knowledge is necessary in BioMedical field
- Quantified Model in ML MD space is a practical approach to describe knowledge and to inspire/create new knowledge

Additional comments

Consideration for Knowledge Founded BioMedicine

- Knowledge Acquisition
    - Communication with users
    - Sharing and collecting thru Web (e.g. Semantic Web)
- Knowledge Arrangement
    - Quality assurance (e.g. EBM)
    - Continuity
- Knowledge Utilization
    - Responsibility and Limitation
    - Applications (Tailor Made Medicine, Drug design, etc.)

Session 4: AI Systems

Setsuo Ohsuga

Waseda University

What is required to AI system?

* AI system as an IT

* IT must be used for helping human activity on informational aspect, i.e. problem solving

* Problem solving is composed of two stages;
 (1) preparation for processing and (2) processing

*Conventional computer technology has been established for the second part under the
 condition that human undertake the first part

* AI systems are expected to automate the first part

* Autonomy in problem understanding, planning, exploring problem solving, etc. is needed

* It is not easy because ; real problems
 (1) concern various types and domains,
 (2) involve human in various ways in the problem

* Objective : To develop AI system

Conceptual Architecture of AI Systems

Problem in the Real World

Knowledge Processing Sub-system

Declarative Language

Problem Solving

Procedural Processing Sub-system

KnowledgeBase

New System User

Problem Domain 1

Procedural Language

Inference Engine

Problem Description

Problem Domain 2

CPU

Program Generation

Problem Model

Problem Domain 3

Program Execution

Program

Programming

Problem Domain 4

Traditional System User

AI System Architecture (Front Half)

Problem

Multi-type, Multi-domain Problems

Network

Externalization Model Building

Problem Model

System Generation
Problem Solving
Decomposition
And Distribution
Functional Integration
Program Generation

Solution

User

Case Base

Knowledge
Base

Knowledge
Discovery
from Text

Text Base

Knowledge
Discovery
from Data

Database

Knowledge Source

Knowledge
Acquisition

Knowledge
Discovery
from Program

Program
Base

Information Source

Information
Gathering

Multi-Strata Object and Multi-Strata Model

Problem Model = Subject Model + Activity Model + Object Model

Subject3 = user
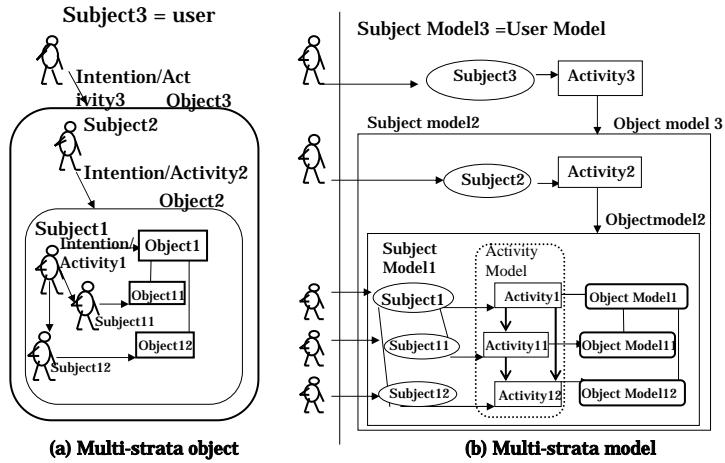
Intention/Activity3   Object3

Subject2

Intention/Activity2

Object2

Subject1

Intention/Activity1   Object1

Object11

Subject11

Object12

Subject12

(a) Multi-strata object

Subject Model3 =User Model

Subject3 → Activity3

Subject model2   Object model 3

Subject2 → Activity2

Objectmodel2

Subject Model1   Activity Model

Subject1 → Activity1 → Object Model1

Subject11 → Activity11 → Object Model11

Subject12 → Activity12 → Object Model12

(b) Multi-strata model

An Example of Problem Model

Primary User

☐ ;Attribute/Functionality
◯ ;Subject
☐ ;Human Activity
⬭ ;Non-human Subject and Activity
▭ ;User Defined Scope

HI   *User Subject*   User's Activity

Subject A   Activity A

UserA   HI

UserB   HI

SubjectB   ActivityB

Object Model

A

UserC   HI   *Sub-C*

C

ActivityC

D

B

UserD   HI   *Sub-D*

ActivityD

Sub-E   ActivityE

UserF   HI

Sub-F   E

Automatic Design

Sub-G   ActivityF

UserG   HI

F

Sub-H   ActivityG

UserH   HI   ActivityH   G

Activity Model 1   Activity Mode 2

Programming ; Activity Model => Program
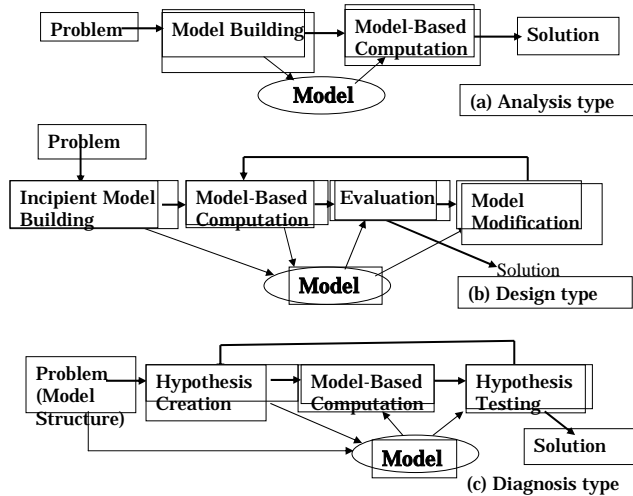
Methodology of Problem Solving (A Few Examples)

Autonomy is achieved by having methodology of doing things

Problem → Model Building → Model-Based Computation → Solution

Model

(a) Analysis type

Problem

Incipient Model Building → Model-Based Computation → Evaluation → Model Modification

Model

Solution

(b) Design type

Problem (Model Structure) → Hypothesis Creation → Model-Based Computation → Hypothesis Testing
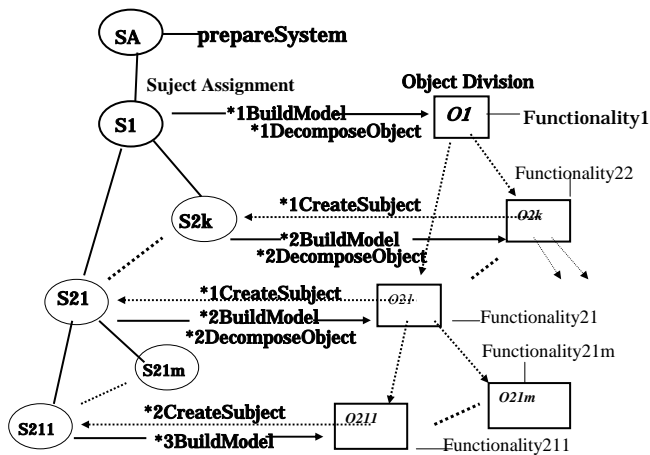
Model

Solution

(c) Diagnosis type

Important Issues for Developing AI System

1 New Modeling Method :
   How are problems represented?
2 Supporting Human Externalization and Model Building - Human Interface:
   How to represent human intension
   How to formalize it?
3 Building Large Knowledge Base and Generation of Problem Specific Problem-Solving System :
   How to manage multi-domain knowledge?
   How to generate problem solving problem solving system?
4 Automatic Problem Decomposition :
   How to decompose problem?
5 Automatic Program Generation
   How to generate program
6 Integrating Heterogeneous Problem Solving Systems :
   How to integrate systems of different paradigm/
7 Knowledge Collection :
   How to find information in Web and assure quality
8 Knowledge Acquisition :
    How to acquire knowledge from data, text and program

Figure 4 An enterprise modeling
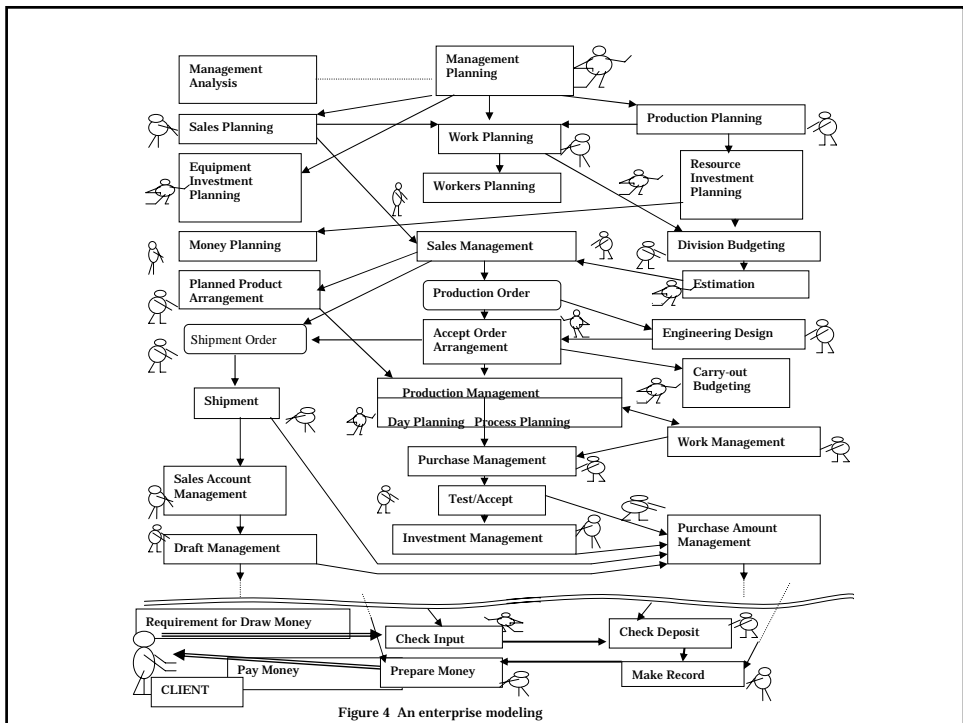
Problem Division based on Object Model Division

O-model decomposition; Object → {Object1, Object2, --- , ObjectN}
  Object-Subject correspondence; | | |
S-model formation;   Subject ← {Subject1, Subject2, --- , SubjectN}

Problem Type and Problem Domain

Problem Type

- **Design**
  - Design-1
  - Design-2
- **Analysis**
  - Analysis-1
  - Analysis-2
- **Diagnosis**
  - Diagnosis-1
  - Diagnosis-2
  - ....
  - ....
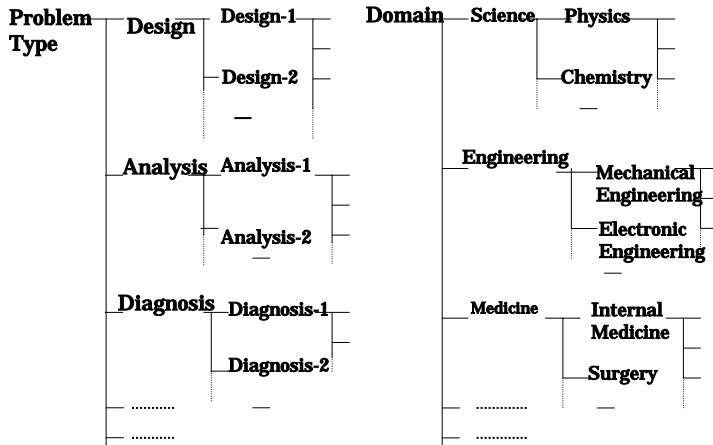
(a) Problem-Type Hierarchy

Domain

- Science
  - Physics
  - Chemistry
- Engineering
  - Mechanical Engineering
  - Electronic Engineering
- Medicine
  - Internal Medicine
  - Surgery
  - ....
  - ....

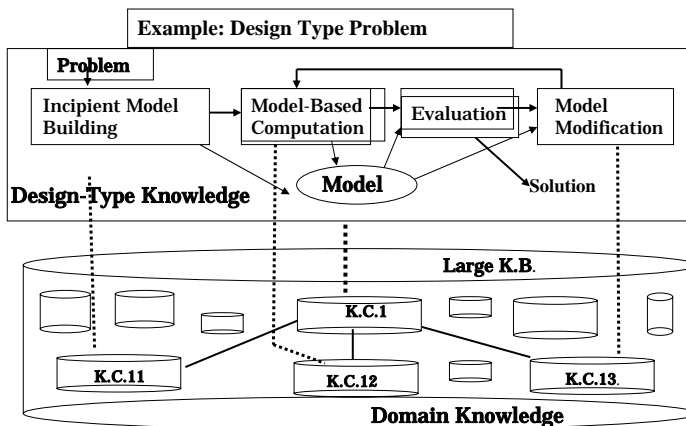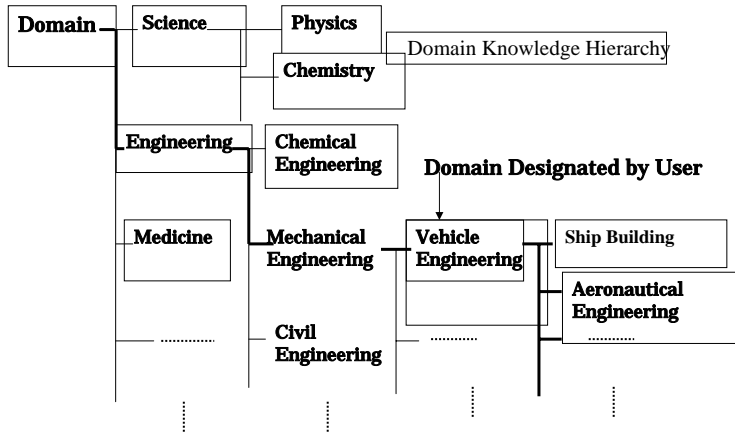**(b) Problem Domain Hierarchy**

Multi-Domain, Multi-Type Knowledge-Base and its Management

Domain Knowledge is combined with problem type knowledge

System; *Extract Relevant Knowledge
         *Construct a Problem Specific Knowledge Base

Example: Design Type Problem

Problem

Incipient Model Building → Model-Based Computation → Evaluation → Model Modification

Model

Design-Type Knowledge

Solution

Large K.B.

K.C.1

K.C.11   K.C.12   K.C.13

Domain Knowledge

109

Knowledge Hierarchy and Identification of Relevant Scope

```
Domain ── Science ──── Physics
                    ┌─ Chemistry          Domain Knowledge Hierarchy
          
          Engineering ── Chemical
                         Engineering       Domain Designated by User
          
          Medicine    ── Mechanical      Vehicle        Ship Building
                         Engineering      Engineering
                                                         Aeronautical
                                                         Engineering
                       ── Civil
                          Engineering
```

Human Interface

Problem Definition : (Problem Type, Problem Domain)

Human Interface :
To guide user to the area in the system defined by
problem type x problem domain

1 To show user knowledge hierarchy with short explanation to each item.
  User points the items (types and domains) that he/she thinks to be close
  to what he/she has in mind
2 Ask user arbitrary set of keywords that the user thinks suited for represent
  the user's intention. System explores the type and domain by special clustering
  method
3 Ask users to represent his/her intension by a simple sentence. System analyze it
  and find items of which the explanation is closest to user's sentence in the
  meaning

Rule for Generating Problem Specific Problem-Solving System

prepareSystem(U-Subject,System):-
 getSubjectReq (U-Subject,U-Activity),
 generateSystem (U-Activity, System),
 evokeSubject (U-Subject, System).

generateSystem (designObject (U-Subject, Model, Domain), System):-
 problemStructure(U-Subject, designObject(HumanSubject, Object, Domain)),
 makeRetrieveKey(U-Subject, Model, Domain,
 designObject (U-Subject, Model, Domain) System), Key),
 retrieveKnowledge(U-Subject, Domain, Key, KnowledgeChunk),
 makeSystem(System, KnowledgeChunk).

evokeSubject (Subj, System) :-
 getSubjectRequirement(Subj, U-Activity),
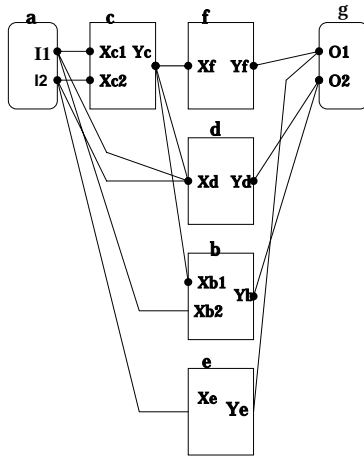 getSolution(U-Activity, System).

retrieveKnowledge(U-Subject, Domain, Key, KnowledgeChunk):-
 getTaskKnowledge(Key, TaskKnowledge),

getDomainKnowledge(Domain, DomainKnowledge),
 includeKnowledge(KnowledgeChunk, TaskKnowledge, DomainKnowledge).

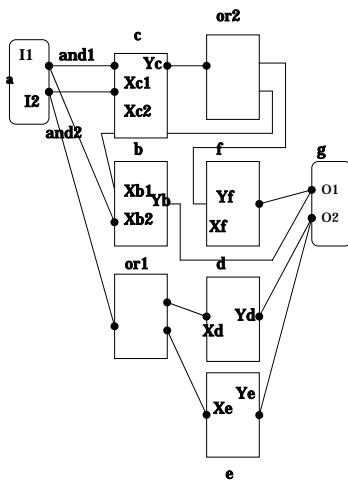
An Application to Experiment of Evolutional System
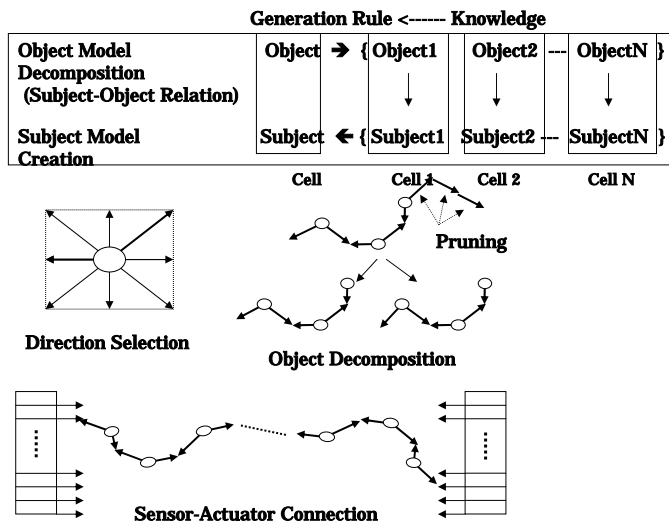
An example of program generation

I1 = [0,5],
I2 = [0, 20],
Xb1 = [6,10],   Yb = [0,0]
Xb2 = [0,5],
Xc1 = [0,10],   Yc = [0,10]
Xc2 = [0,20],
Xd = [0,10],   Yd = [0,0]
Xe -= [11,20],   Ye = [0,0]
Xf = [0,5],   Yf = [0,0]
O1 = [0,0],
O2 = [0,0]


```
Void sol3(int a_a1, int a_a2, int *g_xg1, int
*g_xg2){
        int b_yb;
        int c_yc;
        int d_yd;
        int e_ye
        int f_yf;

        c(a_a1, a_a2, &c_yc);
        if(0<=c_yc && c_yc<=5){
            f(c_yc, &f_yf);
            *g_xg1=f_yf;
        } else{
            b(c_yc, a_a1, &b_yb);
            *g_xg1=b_yb;
        }
        if(0<= a_a2 && a_a2<=10){
            d(a_a2, &d_yd);
            *g_xg2=d_yd;
        } else{
            e(a_a2, &e_ye);
            *g_xg2=e_ye;
        }
    return;
}
```

**(b) One of the structure-of-activities
obtained from the set of the connected
pairs and a generated program**

COMMENTS ON AI SYSTEMS

The Next 5-10 Years

Edward Feigenbaum
Stanford University

AI System Building

- AI's level of maturity as a field requires experimental approach
  - . Need feedback from real world, its complex problems, and issues they bring forth
- AI's models sufficiently complex that to "understand" models system building is needed to perform experiments
- (Congratulations to soccer robotics!)

AI System Integration

- Need to experiment easily with methods of others and complex combinations of our methods and the methods of others.
- Standards-based program interfaces? (APIs for the parts of AI systems?)

Plug-and-Play AI?

- Examples:
  - XYZ Planner?
  - Dietterich's ABC Learning Method?
  - Guha's large KB of people? CYC's KB?
  - Forbus' analogy finder?
- "Best of breed", best practice methods?
- Maybe sources of funding should insist on this form of publication.

AI System Integration (2)

- We may already have more "AI power" than we think we have!
  - But it's scattered in "bits and pieces."
- It would be scientifically useful to see, and perhaps measure, how far we are from human-level AI.

Session 4   Discussions

Paul Cohen:
I want to be careful about our theories. I certainly agree that we might have as much as we need off the shelf to do a lot of really good stuff. In my lab we use off the shelf stuff whenever we can. People are actually discouraged from developing new methods, when old methods suffice. What I'm worried about, though, is that some of us have the sense that we're not winning, that we aren't building artificial intelligence. And when we get together we look around for the usual suspects, like in the end of *Casablanca*. We round up the usual suspects. And among the usual suspects is always architecture. And I don't think that architecture is the reason that we're not winning. I don't think that, for example, the problem that Tom Dietterich introduced us to this morning is an architecture problem. I think that if we had better architectural ideas, then Tom might be able to state his problem in the context of those architectural ideas a little bit, in a way that may be a little easier for other people to see where the problem is. But what Tom said was, "Look. We have to design all of the features, or all of the primitive elements with which our systems are going to reason. There's no autonomy there. We don't have autonomy at that level." That's not an architecture problem. That's a problem of us designing too much for our systems and not requiring a good degree of autonomy of our systems.

And I do worry that you too have an excellent logical argument about using stuff that we've already got, but it might be misinterpreted as saying that architecture is reason that we're not winning.

Ron Brachman:
Firstly, I want to distinguish myself a tiny bit from what I said earlier. What I think I said was—I don't actually believe we have all the piece parts we need. But the way to find out which ones are missing, and what we really need to focus on, is probably to start putting things together in different ways. And then we will ascertain what's missing, what works, and what doesn't work.

With respect to architecture I think you're absolutely right. I was just using architecture to support the systems point, which is if you believe that among the things we should be doing more of should be more integrated systems work, then I just want to make sure. And it's common knowledge that you don't just throw a bunch of procedures together and assume that you're going to get a thinking reasoning system. The architecture matters when you get systems that are this complex. And as I said, the vulnerabilities really tend to be in the interfaces. But, that said, architecture per se I don't think is the crucial problem.

Paul Cohen:
I think the problem is that there is an economic imperative in AI to invest as little as possible in content that's required to perform some task. The engineering end of AI would like to invest almost nothing in content. So, for example, if I want a robot to not crash into a wall the last thing I want to do is teach it about walls. I'm going to tell it, "When your average sonar value has such and such a derivative that passes this threshold, stop." Well I shouldn't be surprised then, that it can't reason about its physical environment because I've invested as little as possible in getting the thing to that level of performance. Moreover, what's in there is something that I put in there. It hasn't been learned. There's no autonomy in the unit. The system doesn't have a clue what

116

those symbols actually mean. All it can do is stop when the sonar value goes below some threshold.

So if you're going to get the kind of autonomous system that you're after, we have to invest in two ways. One is more knowledge. And I think that's universally acknowledged. But also those symbols have got to mean something to the agent, not to us, not only to us. And I think Tom was saying something very much like that this morning.

Tom Dietterich:
In order to try to work on this problem of where the features come from, I would love to be able to pull together some part of the site knowledge base with a learning system and try to re-do the reasoning. So I want that API. And yet, I'm a little concerned from the point of view of research. Research needs to be agile, because when you start out on a project that's not usually where you end up. And so I'm very nervous about going down to build a big system direction when maybe only 10% of that system building effort will actually give a research payoff, especially because I know the appetite of the system integrators in the DOD community.

So I think that the question is how to stay agile and only invest in the parts of the system that are really going to get research value and yet avoid this problem that, when we start to build the system we just hack the parts that we're not interested in and only focus on one bit of it, which also means that we don't learn very much. So—it's a difficult question, sort of research economics. But I think the funding agency is good if they paid for the interface and said, you know, "You're going to be evaluated not by how many journal papers come out of this thing, but by how much and how robust this API is." And that would change a lot.

Koji Sasaki:
Autonomous systems, like AI embedded systems, for example industrial machinery and construction machinery, react to the outside environment like water in the soil. But since the number of variables is limited, it's easy to realize that kind of autonomous system that has already been developed. However, if the amount of knowledge is large, like in biomedicine, it's almost impossible to organize it, so that we have to have a human interact with the system, like a knowledge based system. The point is how to get the good information or knowledge.

Akito Sakurai:
I would like to say my private opinion. We are talking about just the symbolic system. If we bring this theme to, say, fuzzy people, then they will say that symbolic system is not enough to make up a good AI system. The fuzziness is the most important. And if you bring this problem to neural network society, then they will say that this symbolic system is, again, is not powerful enough to build up a good AI system. The connectionist system must be included in it. Is there any opinion about this?

Ron Brachman:
It's all empty words until somebody proves it one way or the other. It's easy to say, "Well, that's not adequate," until the burden is on the folks who are advocating that thing to show that it is adequate. It's nice to have academic debates. That's why I think Ed and I both believe that—I don't think that we're going to push that every single project in AI from now on should be a large-scale systems project. But if in fact you have arguments about the adequacy for some type

117

of technology or another for a task, then it's real simple. You prove it. Let's see if it works. Show me that it's adequate or not.

I think you're absolutely right though, that any line of thinking will tend to denigrate the other lines of thinking and build up their own, and some of the things that we're trying to do is to force people together. I think Stewart has talked about a lot of interesting research that's going on that combines two lines of thought that in his original flow chart showed great diversions at one point in our history. I think that's very promising, very exciting.

Manuela Veloso:
I relate very strongly to the example of the wall and the robot because I face this problem constantly. But the interesting thing is like this. We will only solve Tom Dietterich's problem about learning features if we solve first how to avoid the wall. We won't know what the problem is until we approach it in a way that makes us understand what is actually necessary. So I've learned, during these last few years building these robots that do something from the beginning to end, an enormous amount of things that's hard for me to ever say to someone. It entered my blood, understanding how perception, cognition, and action need to go together. It's useless to have a good vision system if they don't act, if they don't move. The problem with integration—I don't publish papers, it's not publishable material. I cannot say a theory of how to integrate. But, on the other hand, just for us, you cannot believe how much effort goes into a machine, how to operate it. I'm sorry if I cannot make theories about it, but it's indeed a common representation of understanding the level at which things talk to each other, networks of communication. It's true that, sometimes we understand it quickly, but it's something we need to know because we understand things better after we do them. So I really subscribe to the theory that AI needs to be system building at some level.

Stuart Russell:
Systems are our data. We need to study them and then try to figure out why they work.

Daniel Bobrow:
And integration is useless if you are only integrating AI systems. One of things that I took from what Eric was talking about is he had a real system out there. He had the ability to go and monitor it, and look at it. What Koji talked about was about a real system that was looking at integration in the real world. You're going to have to control, you're gong to have to do things with databases, you're going to have to do things with control theory. You're going to have to bring together the kinds of things that Stewart was talking about, and not just from the strands of AI. We have to go back and use the things that are inside computer science and the adjacent fields to actually build these systems. So it's not enough.

Setsuo Ohsuga:
I would like to ask you your opinion on our field and the motivation of our research project. The human society became so complicated that we are faced with very large problems. In this age the humans cannot manage the new scale of problems. Who can solve this problem? If human beings cannot solve this problem, then, it may be necessary to use a different method. It's the reason why we should develop a very large AI system, to deal with very complex problems.

Takashi Washio:

On that issue, I am quite sorry to say, but it's difficult because in human society someone has to take the responsibility for the consequence. If the super-intelligence is to decide something, but nobody takes care of that consequence, then that becomes an issue. Maybe not a technical issue, but we have to also think about the function of the system.

Edward Feigenbaum:

I would like to just briefly defend "Einstein in a box". When I said that, I didn't mean that the box was closed. Einstein did not receive the Nobel Prize for either special relativity or general relativity. He received the Nobel Prize for work he did before that. And, of course, he did one of the first papers in quantum mechanics after that, and so on. So, the Einstein who's in the box is doing remarkable things. It's not doing locomotive things like Manuela's robot. It's not kicking anything or seeing anything, but it's doing remarkable things. Of course I'm not interested in systems that just sit there like a lump of silicon. So it's unfair to criticize Einstein in a box as not doing anything. It's doing thinking. And the question is what more do you want of an AI system? Well, you might want locomotion, you might want vision, or something. But there's nothing wrong with super amounts of thinking.

Daniel Bobrow:

I was going to say that Einstein did not get his Nobel Prize for his thinking. He got it because he wrote papers that articulated to a community in a convincing way the kinds of ideas that he had. You can't think that you're there alone. You have to actually be interactive with the community.

MichaelWitbrock:

I think that he got the Nobel Prize for his thinking. It's not because he wrote a paper for the community, it's the thinking for which he was awarded. I think that perception and locomotion are great, but thinking ought to be enough for now.
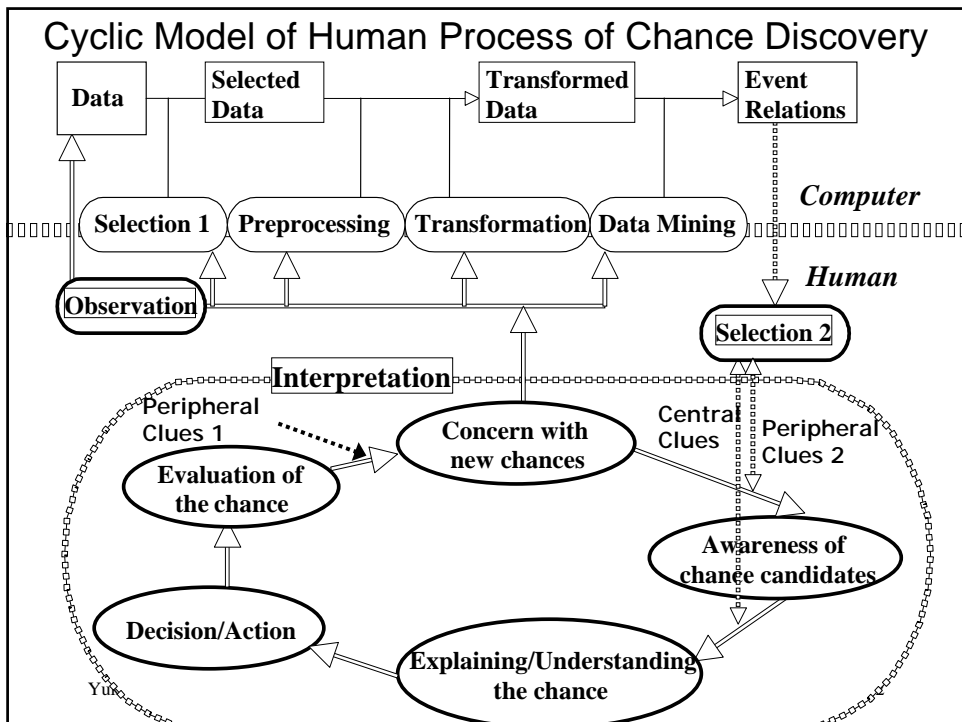
SESSION5: Human Level Intelligence

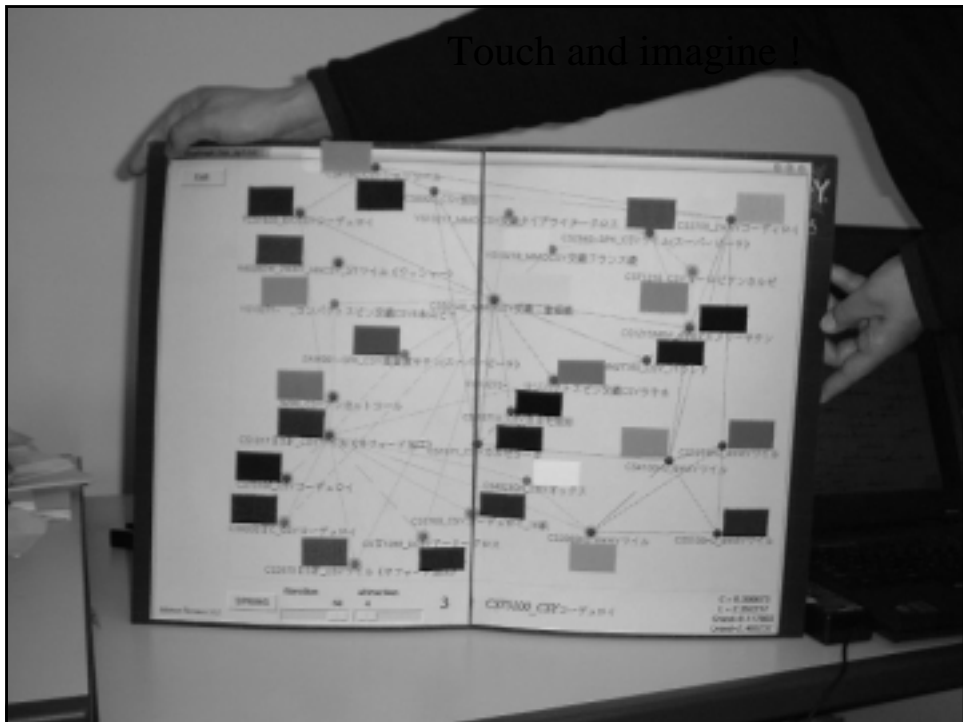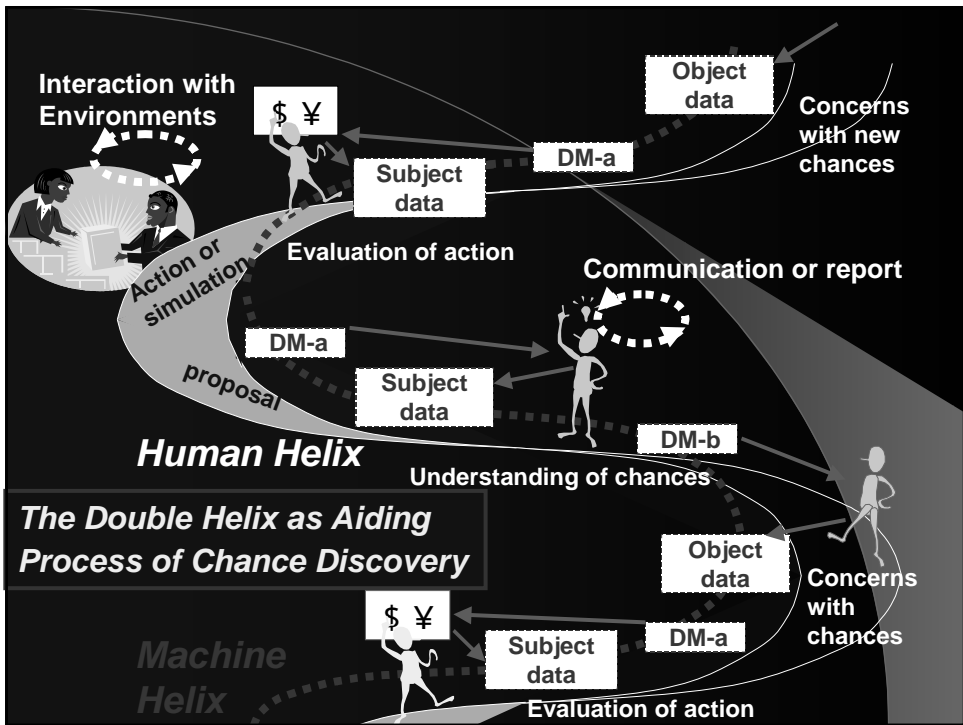"Chance Discovery as Integrated
Human Intelligence"

Yukio Ohsawa
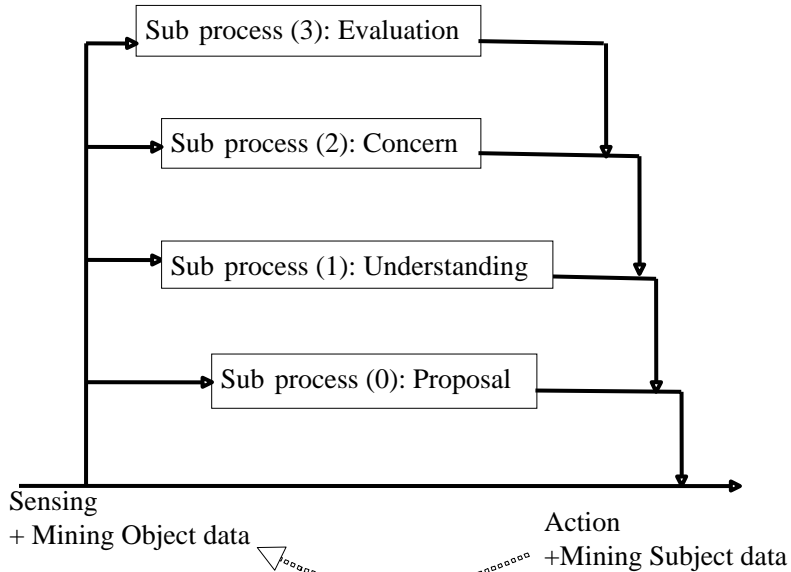
The University of Tsukuba

The Future of AI Workshop

December 14 – 15, 2002

Cyclic Model of Human Process of Chance Discovery

Touch and imagine !

121

To be fast and robust:
Subsumption Architecture of Chance Discovery

Sub process (3): Evaluation

Sub process (2): Concern

Sub process (1): Understanding

Sub process (0): Proposal

Sensing
+ Mining Object data

Action
+Mining Subject data

Conclusion

"Seek, imagine, trust" OR *"trust and seek"* ?

A data mining result says:

-- Japanese people tend to trust human and information which is trustworthy.

-- US people seek, select and trust.

The most Important questions for your chance-discovery are

- **What do you want ?**

- **What do you imagine ?**

- **What/whom do you trust ?**

Session 5 Panel 1 Discussions

Daniel Bobrow:

What is your model for how people use the communication about their explanation? I mean, it looks like noticing something was one aspect, and then starting to articulate it was another, and then hearing how they articulated it helped them transform this. Can you say more about this notion of the transformation of the explanation?

Yukio Osawa:

That's all we have because in the real-world context, communication is very complex. We made some models and we applied some diffusion model. Diffusion model was begun by Rogers in social psychology. That thing could not be applied because the diffusion of information in a community is not so linear as they imagined. Communication is a non-linear matter. One opinion occurs here, and another opinion occurs there, but their mixture makes a quite new opinion. So we do not have any model that can be applied to these kinds of creative complications.

Daniel Bobrow:

When you teach people in your workshops do you teach them about special ways of listening? As opposed to teaching the system, do you teach people differently?

Yukio Osawa:

No. I just show these systems. But now I am in the stage of learning from the users. So the users were interested in their own way of communication. So I had to make a new model for chance discovery complication, anyway. So we do not have any model for that. We can teach, but they teach me.

Manuela Veloso:

How dynamic are these key graphs?

Yukio Osawa:

Key graph is dynamic in two senses. In one sense, when the user does not like the alignment of each node, he can drag it. But also, the key graph is a dynamic in the sense that if the letter is implemented the key graph is revised, according to the implementation. I think you are hitting the point because we are living with the dynamic environment, so the key graph should be dynamic corresponding to the environmental dynamics.

SESSION5: HUMAN-LEVEL INTELLIGENCE

"How should AI assist human situated cognition"

Masaki Suwa

Chukyo University

## The Future of AI Workshop

December 14 – 15, 2002

Situated Cognition

- Cognition does not occur purely internally.
- Cognition and the external world co-develop, affecting each other.
  - ➤ e.g. creative processes,
    learning in a long time frame

e.g. Design process

- Drawing (tentative) ideas on sketches allows designers to detect unintended perceptual features or relations in them.
- That allows for the generation of new ideas or the refinement of previous ideas.
- The subsequent revision of sketches becomes a driving-force for a new detection of unintended features or relations.

Situated cycle
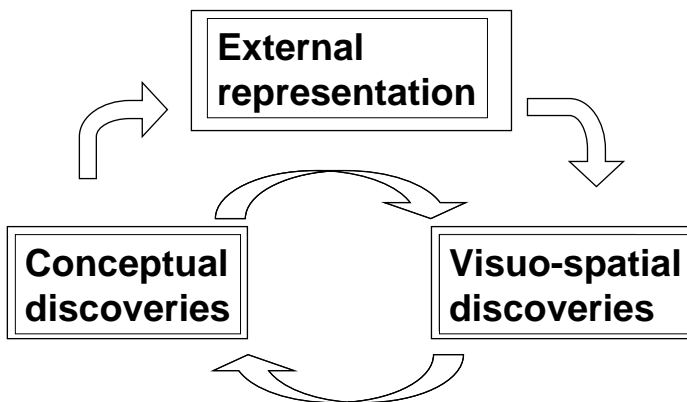
Acts of Problem-finding

- There is no "correct" answer to what should be generated visually or conceptually.
- Situated interactions with the external world suggest local discoveries on the fly, determining the direction of cognition.
- "Goals" in the AI sense are not given, but should be generated out of the situated interactions.

Conclusion

- AI cannot model our situated cognition.
- It should assist our situated cognition locally by
  - showing enriched/multiple visuo-spatial representations for the enhancement of our visual discoveries,
  - providing conceptual cues for us to extend the conceptual discoveries that we make.
- It should not play a role in suggesting candidate discoveries.

Session 5 Panel 2 Discussions

Yukio Osawa:
I didn't say chance discovery process can be modeled by artificial intelligence. I have never said that. I included communication and imagination, and I asserted to him that these things are not model-able.

Koichi Hori:
I still cannot understand why you say you reject another view. To me it seems your view has a lot in common with Osawa's view.

Masaki Suwa:
I'm saying that evaluating is occurring in a very smart and short cycle, so that he may be using some sort of knowledge for evaluating. But the point in the spatial is not the problem of evaluating. Maybe he will evaluate it, but as a result of the evaluating he may have associated some sort of a visual/spatial discovery with the new ideas. And then that may be the driving force for another knowledge for evaluating his own picture of that. He may be able to come up with a new idea. That cycle is important.

Setsuo Ohsuga:
For this calculation process, how do you contact with result?

Masaki Suwa:
In case of design, there is no stopping. That is another point. So maybe the time constraint is a very big constraint for them to stop. As shown in your example, in your question, like where are you stopping, it clearly shows that there is no standard, or there is no legitimate or correct knowledge of when this process should be stopped, and what kind of knowledge should be retrieved, or what sort of ideas should be generated, what sort of visual/spacial discovery should be made. There is no rules and no correct answers.

Setsuo Ohsuga:
I understand what you are trying to say, but the solution should be defined more clearly. But it seems to be a very vague process.

Masaki Suwa:
I don't know what is vague or what is clear. Maybe because of that, it is very difficult to model human behavior in the current AI framework.

Koichi Hori:
I am just curious about whether you reject your own previous studies on the geometric reasoning systems?

Masaki Suwa:
Yes, I think so, because that domain is very related to a logically processable domain, like geometry. And everything can be written in a logic sense, but human behavior is not.

Tom Dietterich:
Let me just say that I think that in design tasks, and particularly architectural design, is a task where many of the constraints and requirements of the design problems are discovered, as you say, interactively, and a lot of it is the process you're describing here, is trying to elicit the requirements and the constraints. So the challenge is that the architect has a vast amount of knowledge about human experience and can mentally imagine parking the car, getting out of the car, walking through the space. And all of these things then evoke other kinds of experiences. And you might say, "Oh, I can hear the noise of the road, I don't want to hear the noise of the road." Or you discover, "Ah, there's an opportunity with the road," and so on. Now, I think that this means that the economics of building an artificial intelligence system that has all of that knowledge is very challenging because it's a huge amount of knowledge to represent. And the human already has all of that knowledge. So perhaps the value of the computer in this case is more to try to record the discoveries that the human makes and not lose them, because a problem, often, in design is that you have an idea but it gets lost in a series of meetings and so on over time. So design, particularly creative, architectural design I think is one of the most difficult of areas for the economics of AI. But I don't think that it's impossible to do. If you think about people who design integrated circuits, for example, these systems have a vast amount of knowledge about the design constraints, and they can check the uppermost design and simulate in vast detail what's going to happen. So this is a case where the computer supports the design process and really is better than the human designer.

Masaki Suwa:
Maybe the point is, as you say, the amount of knowledge that has to be incorporated in the computer. When I think about the amount of knowledge which should be put into the computer in order to simulate, for example, to simulate the architect's behavior, I am very pessimistic because as I look into the quality of processing of the architect more and more, I don't know what is the boundary of the knowledge they use. Sometimes they may be talking about the function, and in other cases sometimes aesthetics, and sometimes all human experience, like psychological states or something.

Stuart Russell:
The architecture department at Berkley has several faculty members who are developing artificial human agents that they can put into simulated buildings to find out whether those buildings work. And they often find that, no, the building doesn't work and they have to fix the design. And it's one of these 80/20 things. The first few times you do this you find that by putting in a few, simple bits of knowledge about humans you can discover a lot of major flaws in buildings. And then you put in the next 20 things and you will discover a few more flaws about the building. And maybe you will never get something that perfectly simulates a human, but with a reasonable amount of effort you can get a reasonable return from that process. I don't think there are reasons for pessimism.

Masaki Suwa:
Maybe. I might be saying that I am pessimistic when I think about the very creative processes. So reasonable processes and creative processes can be divided. And I may be showing a model of a very creative process.

Daniel Bobrow:
"Impossible" takes a very long time to prove. And I think part of what you're doing is trying to compare where we are now. And one can say, "Well, if I have I. M. Pei doing the same thing, it might be different than if I have Frank Geary doing the same thing. And Pei will not see the same things that Geary does." And we might be able to take advantage of the differences of perception of what an AI might do, and it might not see the same thing. It might not replace a particular human being, but may be able to suggest things, which in connection with other humans talking about the design, we'll be able to see some new relationships. I think that's part of what Stuart was saying, so that—I think it's very important to distinguish exact match with specific individuals and the ability to get things which we might think are creative—seeing new relationships that come out of, that emerge, from the plans. And those suggestions could be often useful to other human beings, as well.

SESSION5: HUMAN-LEVEL INTELLIGENCE
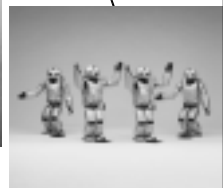
"Robots as Intelligent Beings"

Manuela Veloso

Carnegie Mellon University

The Future of AI Workshop

December 14 – 15, 2002

Robot Tasks

- High risk – "super being"
  - Space, medical robotics, driving, factories, rescue, underwater, military

- Low risk – "regular being"
  - Office assistant, museum guide, soccer player, coach, nursebot, entertainer, …

Thanks to Sony, CMU illah, thrun, reids

Manuela Veloso   The Future of AI Workshop

Autonomous Intelligent Robots

- Perception
  - sensing the physical world, multi-sensor fusion
- Cognition
  - action selection, reactive, deliberative, planning
  - interaction with humans, and other robots
  - reflexion, adaptation, learning, collective modeling
  - multi-robot coordination, teamwork, multi-agent learning
- Action
  - motion, navigation, obstacle avoidance , manipulation, dialog, gesture, expression, emotion

Coexistence of robots and humans in natural environments:
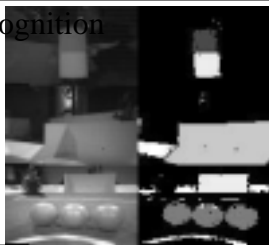***THE* challenging task.**

Purposeful Perception:
Reasoning as Bias Provider and User

• Object Recognition

- modeling
- real-time detection
- automated calibration

• State-Action Machine

not see ball → Recover ← timeout

Score   see ball   Search

next to ball   not see ball

not next to ball → Approach ← see ball
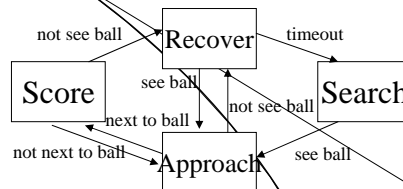
• Landmark-Based Localization
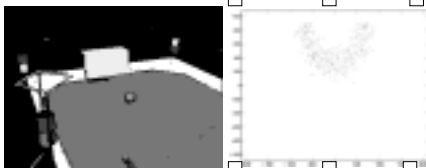
• Behavior and Plan Mining

acce

S_1 →(o1) S_2 →(o1,o2) S_3 (o4)

o1,o3 S_4

reje

$$\hat{S}_t = \arg\max_{s_i} P(S_t = s_i | o_1 ... o_t)$$
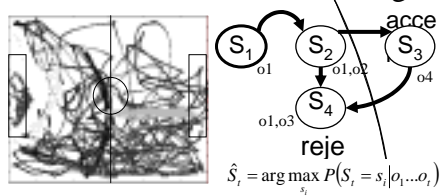
Multi-Robot Planning, Coordination, and Learning
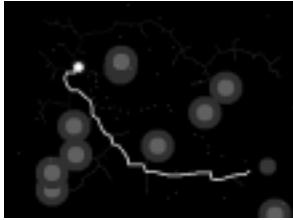
• Real-time path planning with obstacle avoidance

• Multi-robot planning and real-time adaptation

• Dynamic gradient-based team coordination

• Multiagent control learning; convergence, rationality, limitations

Role 0
• Dribble to P_1
• Pass to R_2
• Wait for loose ball
Role 1
• Wait for Pass at P_2
• Receive Pass
• Shoot

$$\mathrm{P}\!\left(p^* = p_j\right) = \frac{w_j}{\sum_{P_A} w_k}, \forall j \in P_A$$
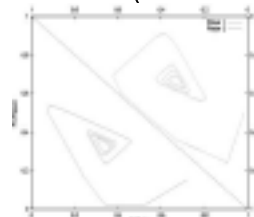
Manuela Veloso   The Future of AI Workshop

Conclusion:
A View of Robots as *Beings*

● Robots have *complete* intelligence
– Perception, cognition, and action
● Robots have limitations
– Failure, success, learning
● Robots coexist with humans
– Assistance, interaction
● Robots extend humans in *some* dimensions.

Thanks to M. Bowling, B. Browning, J. Bruce, S. Lenser, D.Vail, E. Winner
http://www.cs.cmu.edu/~coral

Manuela Veloso   The Future of AI Workshop

132

Session 5 Panel 3 Discussion

Michael Witbrock:
In answering these reporters, you say that robots extend humans in some dimensions. But, you know, humans extend cats in some dimensions. It's not a very satisfying reply. We've definitely taken over from the cats.

Manuela Veloso:
I'm really bad about waving my hands about things I don't know. So when I'm faced with these things about ethics, and robots, and all of that I just say, "Who cares?" I am passionate about my robots. I will continue doing that. But I guarantee to you that my belief that they will be omnipotent chess players, soccer players, excellent parsers, natural languages, and all of these all in the same picture at their maximum—it might be very improbable because, you know, in some sense these little soccer guys only do soccer. They don't do anything else. There was a little girl who asked me like this, "Do the robots wonder why people pick them up?" This is a seven-year-old kid. And I was giving this talk at the science museum. And I said, "Any questions?" And she said, "Do they wonder why people pick them up?" She didn't ask why do you pick them up? "Do they wonder?" And I looked at my student, and I said, "no." And you have to realize that it's so—it's very deep.

Another thing that happens with children, or with questions I get, is that you scream and shout at these robots, "Shoot! Go!" You name it. And you know what? They can't hear anything. They have sound processors, but not speech processors. So I've explained this to technical people and also children. And I'm telling you now they don't have a microphone. They don't have anything. At the end of my explanation, being very technical, very honest about what I do, this kid comes to me and said, "They don't hear you, but they hear me." And I said, "What do you mean?" "Every time I said 'Shoot!' they did."

And I'm like saying, "I'm guaranteeing to you they cannot listen." And he went away absolutely convinced, or maybe convinced at least, I hope, that indeed these creatures could listen to him and not to me even if I show there are no microphones in the little thing. So what I'm trying to say is that robots, they only do soccer, that's true. But look at how a little dance makes you think, and makes people happy. Now I got bored because they did this dance 30 times, because every time they scored they were doing the same dance. But for 2003, I'm going to have ten different dances. And I'm going to have a random selector of the dance. And I'm going to be surprised at the dance they do. And I'm going to have fun, too. Because after they score, do they all go and get together in a circle, do they do like that, do they do like this? I don't know. So we are all going to say, "What are they going to do?"

Michael Witbrock:
Perhaps this is a use for the microphone. You can just measure the sound level and use reinforcement learning on the dancing.

Eric Horvitz:
When you refine your methods you rely upon a set of heuristics about what you can do in a hill-climbing or hill-descending manner. And you rely on harnessing local sensing and deliberate

about action within a short-term proximity. I am curious—looking at your gradients, for example—what is it about robotics or about this particular set of tasks and context in robotics that tells us something about the utility of myopia and the ability to depend on assumptions of locality in the spatial/temporal reasoning about games of soccer. And how might the results here map to other kinds of things that you may be working on.

Manuela Veloso:
That's very interesting. That also goes back to what Ed was saying yesterday. We have been making a lot of effort to extract modules that can be applied independently from these steps. The one that we have been most successful on pulling out has been the vision. So there is now a module called CM Vision that's used by even factories or other people. That is color processing in real time. That is object recognition. The gradients that we are doing are also very general in terms of you giving a traction and repulsion points, and giving the right functions that they should use.

Eric Horvitz:
But we can't always assume that such smoothness exists in our problem solving.

Manuela Veloso:
That's true. I can only tell you that if you can identify the wall as traction and repulsion points, you can actually try to apply these types of gradients to the work. But if it's not, then we have the assumptions of the work. I'm not using any of the position, yet, of the opponents, for example, because I don't know where they are very well. The amount of noise of the opponents doesn't help my gradient. So I only use where the ball is and my teammates, because they talk between each other. But you are interested in that they don't spread beautifully as where the others are. It's a challenge. And then, it's a challenge to go from a system that really does something impressive at the AI level to the actual feelings about how does it work, and why is it doing in the mind. And I guarantee you that I'm committed to try to address that challenge. But on the other hand, I want action for the system, too. So I'm going to always try to make the system always better even if I cannot expect immediately the feelings in their mind. I just go slowly.

SESSION5:

Human-Level Intelligence

"Making Humans More Intelligent through Collaboration"

Naomi Miyake

Chukyo University

The Future of AI Workshop

December 14 – 15, 2002

Why Bother?

- "Psychology must be the master science in the very distant future (to offer an effective control system of the human creativity)… from now on social systems may fail, not because of corruption or defeat in conflicts, but because of their inefficiency as systems in an extremely energized society…we need to find entirely new dimensions for systems to absorb a lot of new energy, and such a new dimension may be found only through knowing man better. "

Toda, M. (1971) "The role of psychology in the very distant future" at XIXth ICP.

We Need
Better Cognitive Science

- We need better models of human cognition, both for promoting the master science a la Toda and for educating people so that they can take better control of themselves.
  - *Toda is a great scientist, but has never made a good teacher...*
- To do this, collaboration may be the key.

Why Collaboration?

- Because both intelligence and science are fundamentally social processes.
- Collaboration has been shown to be effective in finding new ideas in scientific research.
- Collaboration has been found to provide students with rich learning environments for knowledge building and reflection.

Technology Can Support Collaborative Model Building

- Success case: "Can we have ice everyday?"
  - "Let's leave water in a container you choose and see if it freezes." --- The experiment lasted for two weeks and children got rudimentary understanding of what causes water freeze.

- Technology can help raising
  - Visibility/share-ability of solution variations
  - Tracking capability of integration efforts

Conclusion

- We need better models of human creativity and human learning, a lot more strongly now than before.
- Collaboration may be the key for this model building and learning, both for scientists and for lay people like college students.
  - We need to know how to teach cognitive science broadly.
- Technology can help us know ourselves better, by providing ways to externalize and keep better records of our own cognitive processes.
- AI is welcome to collaborate with us humans.

Session 5 Panel 4   Discussion

Kumiyo Nakakoji:
This is a question not only to Naomi-san, but also to the whole panel, that there seems to be two dimensions, or two types of levels for human intelligence. One is sensory modeling and processing level in a perceptual and conceptual way. And the other dimension is individual versus collaboration in a social setting. And it is interesting that from computer's or robot's point of view there is no difference, model sensing versus collaboration, because that is yet another model sensing scheme. So I'm wondering if it's a single dimension, like layers and layers starting from sensory motor skill to the brain, and then perception, and then conception model, and somehow the social communication and relationship come in. Or is it a totally different thing for humans? But the success of the robot case, coordinating everything and this purposeful perception, it seems to be some indication that they can be the same thing.

Eric Horvitz:
There has been recent research, some on our team, in the realm of automated cinematography and media capture where investigators have tackled challenges with summarization, distilling from large streams of data and experiences what we might refer to as the important "key frames of life" and organizing them in a way that's useful given the expected goals associated with future review. I could see the tools you're using providing value in educational settings in a number of ways. Do you have comments on the AI challenges in automatically producing these graphs and so on?

Naomi Miyake:
I have a very simple question to add. Humans use a different sense to make a summary version of a video. And there has been a lot of debate amongst teachers and teacher education people that who would be doing those clipping, and summarizing, and showing those long classroom videos in a five minute gist to show the best part of the teaching. And to me, we probably do not have to decide or select one of many. Because of this capability, it's not like somebody just makes a clip for summarization, but summarization should be accompanied by explanation, or the claim of what viewpoint that those summaries were made from. And AI, with its capability could come and join us on making what kind of reasons that those systems made those summaries with, juxtaposed with human-side summaries. And now we can see that there are different ways to see the same classroom and the same video, and that's the power that we can have. It's not the power of some algorithm winning over the other, but having variations of things and showing those variations, giving us some chance to compare and think about those things.

Kumiyo Nakakoji:
To follow up on your question, and also coming back to my original question, is that Suwa-san's talk clearly indicates that doing is a source of understanding. So in her application domains, like cutting and identifying which party is important for me, and for you, and for them, if that is a part of the learning process, automating that part is in some sense depriving us of the opportunity for learning.

Yukio Osawa:

This is my answer to Nakakoji-san's question about the two dimensions—one is perception, and the other is communication or social vision. I think these two things seem to be independent of each other. It depends on what you perceive. If the target of perception is in the past or in the current situation, I think these two things are different. But if you want to make a perception of future events, it includes various other factors, including imagination. And this imagination is triggered by the memories of one's past experience. So here, analogical matching of the current situation and the past situation triggers the imagination for the future situation. And this analogy is what humans or maybe robots, have to communicate with others. So in this sense, communication can be a sort of trigger to perception of the future.

Naomi Miyke:

Suwa-san, you don't use very many collaborative situations, but do you have something to say, whether the individualistic design process would be different from designers working together, or designers working with a system, or in what fashion would they be different?

Masaki Suwa:

My investigation is just looking at the individual, but I'm not claiming that collaboration is not necessary. I think that it is necessary. And maybe the reason why I have been looking at the sketching is the same reason as you are looking at collaboration. Collaboration and sketches are the same thing if you think that both represent an external world that's surrounding you.

Manuela Veloso:

This is where I have a big advantage working with robots rather than with people. I actually can decide what they tell each other at the design of the robot. And I'm limited by bandwidth, by real-time aspects. And one day, if we have infinite bandwidth, I could have a robot share with another robot all of it's life, all of it's perception. Whether it is useful or not, I don't know. But if you want me to tell you all of my life, even if I want, I can't, through communication. So when we collaborate we are limited by our own cognitive processing. The robots are limited by bandwidth, by time, but I could make them talk about everything. We obviously disagree on many things, but robots disagree tremendously about what they perceive, what they think the world is like, their models. They move in the world seeing the ball all at different places with the errors of 5 or 10 centimeters. Because of the errors in sensors and because of the occlusion, and because of the walls, the lighting and everything, it's impossible. So in some sense, sometimes you feel like turning off communication. "Leave me alone and let me do it just according to what I believe is true." So there is, research-wise, a big problem between the opportunities that multi or collaboration offer versus actually doing it all just by yourself. Intuitively we all believe that collaboration is a good thing. Mathematically, it's a nightmare. So we have to try to solve this communication part at the algorithm level better. And that's why, for example, the team we played against in the final did choose explicitly as their design not to have the robots communicate at all. Well, they lost because they could not see the ball across the field.

They beat us 9-1 the year before, or something. So they could have easily won again because their motion is faster, if they just would see the ball more often. They spend their life, their time, searching for the ball. If they had communication between the robots like we did—we were slower that they were. I did not know that we were going to win at all. I was just as surprised as

139

they were by noticing the fact that we could share the perceptions made such a big difference. It compensated for the speed at which they moved, much faster.

SESSION6:　KNOWLEDGE PROCESSING

"Digital Content Annotation and Transcoding
for Web-Scale Knowledge Systems"

Katashi Nagao

Nagoya University

The Future of AI Workshop
December 14 – 15, 2002

---

1. Towards Advanced Sharing and Reuse of Knowledge

- *Annotating Digital Content with its Semantic Descriptions*
- *Customizing Digital Content according to User Preferences*
- *Discovering Knowledge from Semantically-Annotated Digital Content*

2. Web Superstructure

- *Current Web consists of hyperlinked documents.*
- *Annotations are meta-level content and constitute a hierarchical structure.*
- *Web content and its annotations make a superstructure on the Web.*
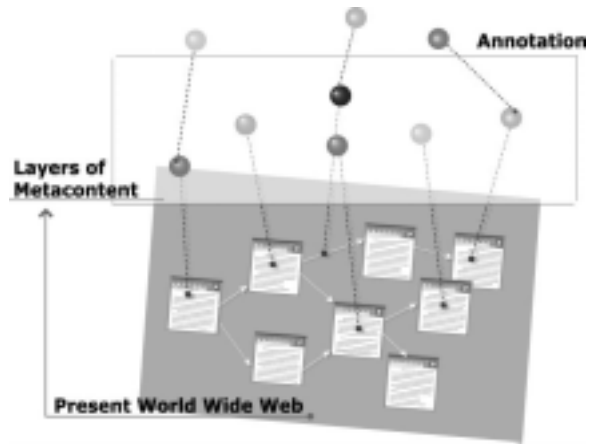
3. Semantic Annotation

- *Annotating Digital Content with its Semantic Descriptions*
- *GDA by K. Hasida*

Annotation Editor


```
<su id="id1-0" ptb="S">
 <ajp ptb="DT">the</ajp>
 <n bf="power" ptb="NN">power</n>
 <adp>
  <ad ptb="TO">to</ad>
  <vp>
   <v bf="make" ptb="VB">make</v>
   <np ptb="NNS">decisions</np>
  </vp>
 </adp>
.</su>
```

Example of an annotated text

4. Semantic Transcoding

- *Customizing (Annotated) Digital Content according to User Preferences*
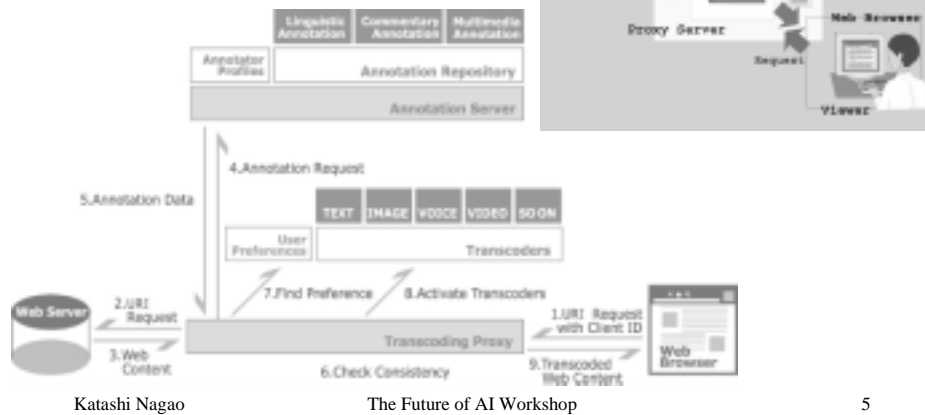
5. Conclusion

- *A very important thing is remaining.*
  - *That is "Knowledge Discovery from Semantically-Annotated Content."*
- *We are also developing a technique for semantic retrieval and on-demand editing of content including multimedia data.*
- *Relations with the Semantic Web ...*

Session 6 Panel 1   Discussion

Einoshin Suzuki:
Can you show an example of a semantically annotated document?

Katashi Nagao:
This is an example of a text with tags using XML format. But this is not manually made. This was automatically created using tools.

SESSION 6: <u>Knowledge Processing</u>

"Using the Ontology to Grow Itself"

Michael Witbrock

Cycorp, Inc.

The Future of AI Workshop

December 14 – 15, 2002

A Large Knowledge Base

Cyc contains:
- 10,000 **Predicates**
- 100,000 **Concepts**
- 1,400,000 **Assertions**

Represented in:
- **First Order Logic**
- **Higher Order Logic**
- **Micro-theories**

<u>Connects</u>
Web, DAML
DB's

<u>Construction</u>
100's person years

Thing
Intangible Thing — Individual
Sets Relations — Spatial Thing — Temporal Thing
Space — Paths — Partially Tangible Thing — Time
Spatial Paths — Logic Math — Events Scripts
Physical Objects — Borders Geometry — Artifacts — Agents
Living Things — Materials Parts Statics — Movement — Actors Actions — Organ-ization
Ecology — Life Forms — State Change Dynamics — Plans Goals — Organizational Actions — Types of Organizations
Natural Geography — Human Beings — Plants — Physical Agents — Organizational Plans — Human Activities — Human Organizations
Political Geography — Human Artifacts — Human Anatomy & Physiology — Animals — Agent Organizations — Business & Commerce — Politics Warfare — Nations Governments Geo-Politics
Weather — Products Devices — Conceptual Works — Emotion Perception Belief — Human Behavior & Actions — Social Behavior — Purchasing Shopping — Sports Recreation Entertainment — Professions Occupations — Law
Earth & Solar System — Vehicles Buildings Weapons — Mechanical & Electrical Devices — Software Literature Works of Art — Language — Social Relations Culture — Social Activities — Transportation & Logistics — Travel Communication — Everyday Living — Business Military Organizations

Domain-Specific Knowledge

(e.g., Bio-Warfare, Terrorism, Computer Security, Military Tactics, Command & Control, Health Care, ...)

Domain-Specific Facts and Data

What's It Good For?

- Basis for Machine Learning
- Observe Regularities in Assertions
- *Interviews:* Seek Similar Knowledge

(isa GeorgeWBush Politician), (PolticalPartyMembers GeorgeWBush RepublicanParty)

(isa HelenClark Politician), (PoliticalPartyMembers HelenClark NZLabourParty)

(isa JunichiroKoizumi Politician),(PoliticalPartyMembers JunichiroKoizumi JapanLDP)

Rule induced:

(isa ?pol Politician) -> (thereExists ?par (PoliticalPartyMembers ?pol ?par))

Knowledge Entry Requiement added:

(isa ?pol Politician) -> (keStrongSuggestion ?pol (thereExists ?par (& (isa ?par PoliticalParty) (PoliticalPartyMembers ?pol ?par))))

Making Rules do More Work

Pre-existing Rule: For every EVENT and JOB-TYPE, if something's occupation is JOB-TYPE and *EVENT is a duty related to the job of JOB-TYPE*, then the duties of an individual who has JOB-TYPE include EVENT.

New Fact Added to KB: Yoriko Kawaguchi's occupation is foreign minister.

Cyc Asks: Is it true that there is some EVENT such that EVENT is a duty related to the job of foreign minister?

- Seek to satisfy antecedents of existing rules
- Maximise inferential power of new knowledge
- New facts can form basis of rule induction

Automated Question Answering

Current Research:

- New Knowledge added leads to interview
- Cyc attempts automatic answer using IR, IE, parsing retrieved text, e.g:

Interview Rule: If something is a kind of animal, then Cyc should try to find out things that it will eat.

New Fact: Pigs are members of family Suidae.

Cyc Asks Itself: What do pigs eat?

IR Query: In phrase: Pigs, eat or consume

Retrieve: "*Experts say that pigs can eat crushed stalks* after it has fermented with a kind of additive…" from TREC corpus

Dependency Parse and Add Answer to KB: Pigs eat stalks

Conclusions

- Effective learning requires inductive bias
- Large knowledge bases provide one
- Helps less-trained people add knowledge
- Allows automated knowledge addition

Next few years:

- Further automate into self sustaining process

Try Cyc at www.opencyc.org

Session 6 Panel 2   Discussion

Tom Dietterich:
You talked a little about how you evaluated, but can you say a little bit more about evaluation of the knowledge that it discovers for itself?

Michael Witbrock:
This is embedded in our regular knowledge formation tool. It tells the user what it has inferred as the knowledge seeking goal. It tells them what instances it inferred it using. It gives the reasoning used to do that inference, and gives them the opportunity to say that this is a silly question. So every time one of those events happens, Cyc sends off e-mail to us telling us that it was embarrassed by what we have taught it to do. So if you're actually asking if we have rigorous evidence, no, I don't know. We haven't done a careful experiment yet. So anecdotally, I'd say about half of the things that it decides to ask you are really good questions, and about a quarter of them are ludicrous, and about a quarter of them don't make any sense because they involve mysterious philosophical concepts, which we haven't yet discouraged it from telling you about.

Ron Brachman:
Michael, imagine looking way down the road. And you get all the people in the world to contribute to Cyc, and you have billions and billions of assertions. What else is left to do? That is, have you finished AI in the sense that we were discussing it yesterday, or does that provide one tenth or one half of the whole?

Michael Witbrock:
Well, billions and billions of assertions don't give you all that much. They give you billions and billions of assertions. One thing that we at least need are inference engines fast enough to make billions and billions of assertions useful. And we don't have that yet.

On top of that, of course, it's important to be able to do things like vision. It's not clear how to do those just by assertions and first order predicate calculus. And if you could do it, you'd need a really fast inference engine to support it. So I think there's still work to be done. So there is any amount of research that is possible to do. For a start you want to do rule induction. And we really want to use reinforcement learning to speed up learning. We think they can give us enormous boosts in performance.

So there are opportunities everywhere in using machine learning. And I'd like to see a large push in that direction over the next few years. I think that's going to make this process of ontology and producing knowledge worthwhile. Just producing ontologies—and some people are excited by that. I couldn't care less about them. I want to produce an AI that uses ontologies.

Manuela Veloso:
Do these possible billions include procedural knowledge, like recipes, how to cook a good omelet, and how to change a tire, and how to solve a second degree equation? If so, will Cyc also have all the Dijkstra's algorithm all the way to cooking an omelet?

Michael Witbrock:

Currently there are sets of algorithms for a number of things that we have detailed scripts of. We're certainly moving in the direction of representing more and more procedural knowledge. We are moving in the direction of explicitly representing the behavior of things like our knowledge formation tool, where we believe that it's the only way to make acquiring knowledge of different types in the general sense truly productive. At the moment it does the same sort of thing for every kind of knowledge. That makes no sense. It should be driven off a script, preferably off induced scripts so that it if you are telling the system about, for example, computer scientists, it asks you the right questions in the right order.

So that's trying to maximize its information gain in the direction of things which are useful to know about computer scientists. So they're looking, for example, if it can tell you about a conference. It should give you a table for the conference participants rather than asking about each of them individually.

Stuart Russell:

One of the things you mentioned is the ability to suck in databases and other external sources of information. But that seems to bring up the naming problem in space. You may have Koizumi in your knowledge base, but then you may have a telephone book database where Koizumi is identified by a social security number or something like that. How are you going to solve this? It's not an easy problem. I don't think it's insoluble, but it's not easy. For example, when you have 100,000 names in your system, and a database has 27 million names. And there's no connection between those two sets of data.

Michael Witbrock:

Well, there is a connection. When you are describing a database to Cyc you tell it what the tables mean in Cyc terms. Also, there's a different question of how you work out whether someone with a name is the same person as someone else with the same name or not. And that's sort of a general problem.

Stuart Russell:

Yes, I agree that this is a general problem. The database people have been stuck on this. I mean, they essentially do it by hand.

Michael Witbrock:

Well, one way of fixing this is if you can't help by automated means, for example, then you can certainly make rules that say if someone has an especially unusual name, and they live somewhere, and in the database it says that they live somewhere where they're actually believed by Cyc to live, or Cyc already knows their social security number that corresponds with the social security number in the database, then you can infer that those names are the same person. And I guess the only solution to that is to give the system as many rules like that as you can come up with. And with a knowledge base you're certainly in a better position than someone trying to do databases, because you have a very sophisticated vocabulary for describing those rules. But in general, it's an unsolvable problem. People can't solve this problem, in general. Ultimately, we resort to asking someone whether they're the same person. And ultimately Cyc will have to resort to asking someone.

Koiti Hasida:
Are you trying to make a huge concept tree for everybody, or a collection of small trees for a small number of people?

Michael Witbrock:
Sort of both. We're trying to make a huge knowledge base which is capable of representing all human knowledge. However, inside that knowledge base we have partitions which form particular sorts of tasks for different types of reasoning. A system can decide to reason only in a partition. For example, we have a partition called "chemistry," a chemistry micro-theory. We've got a project at the moment to have Cyc learn how to do high school chemistry exams. Someone's funding us to do that. And most of the reasoning for that particular type of problem will go on in a very small part of the tree. But when the system has to infer, for example, that something is a substance, then we'll use the representation of substance in the main knowledge base.

Koiti Hasida:
So on small trees, a small number of people can agree for that tree. But for huge trees, I think it's difficult for everybody to agree.

Michael Witbrock:
People don't have to agree to the representation in the knowledge base. What is necessary is that people will be able to communicate with the knowledge base. For the ontology, it's not important for you and me to share exactly the same representation of the concepts. It's not important for me and Cyc to share exactly the same representation of the concepts. What's important is for Cyc's representations of concepts to be translatable into a form that I can understand and vice versa. So I regard the natural language interface to these knowledge bases as absolutely essential. In some sense they are almost meaningless without a translation into a natural language. And we give concepts in the knowledge base meaningful-looking names, but those names are in some sense a fraud. It would be more honest to name these things by hexadecimal numbers. Then by looking at them you would really understand what the knowledge base is, because that's all it is to a computer. It's a bunch of undistinguished symbols, which it processes. So at the level of the knowledge base that is what there is there, and only by converting that either to natural language, or in fact to sensors and data on robots, can we give them meaning.

SESSION6: KNOWLEDGE PROCESSING

"Annotation-Based Grounding"
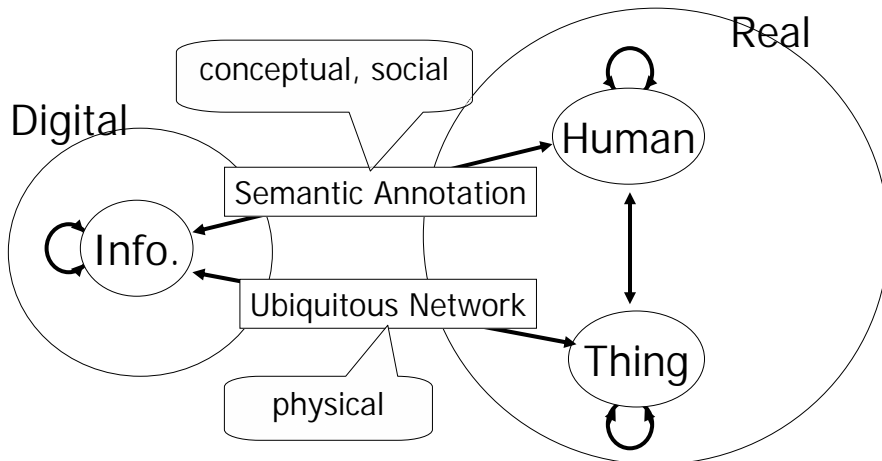**Koiti Hasida**

CARC, AIST & CREST, JST

The Future of AI Workshop
December 14 – 15, 2002

Grounding

- To endow digital information with inherent real meaning
- Traditional approaches
  - Pattern Recognition
    - vision, sound, language, etc.
  - Embodiment
    - robot
- Premature technology

Collaborative Grounding by Information Infrastructure

Real

conceptual, social

Digital

Human

Semantic Annotation

Info.

Ubiquitous Network

physical

Thing

GDA (Global Document Annotation)

Interpretation of propositional content
   retrieval, summarization, translation, etc.

```
<su>
  <np opr="agt">Tom </np>
  <v sem="past.eng:meet">met </v>
  <np>
    <adp sem="sg">a </adp>
    <n opr="obj" sem="eng:girl">
      girl
    </n>
  </np>.
</su>
```
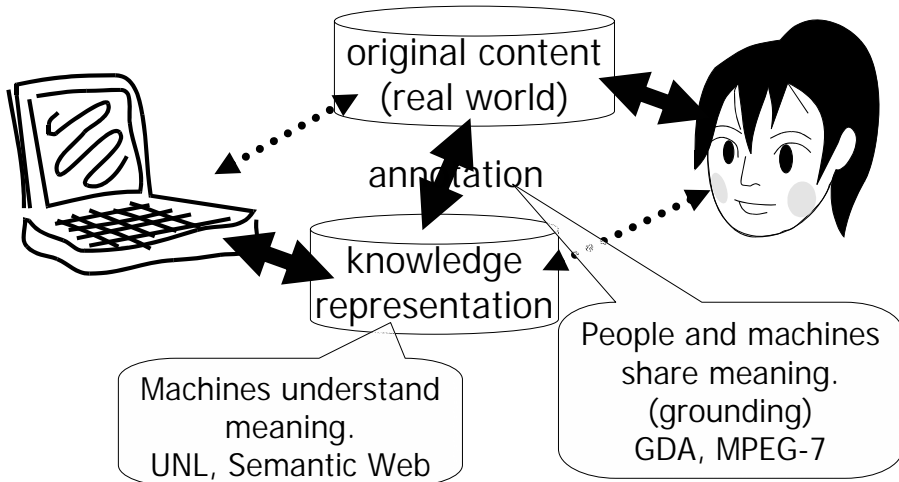
Annotation　　Grounding

original content
(real world)

annotation

knowledge
representation

Machines understand
meaning.
UNL, Semantic Web

People and machines
share meaning.
(grounding)
GDA, MPEG-7

Concluding Remarks

- Social information infrastructure for grounding
  - annotation-based intelligent content
  - location sensors, mobile communication devices, etc.
- Semantic annotation
  - conceptual/social grounding
  - propositional content
  - interaction-based grounding & grounding-based
    interaction
  - research data & end-user content
  - from Semantic Web to Semantic World

Session 6 Panel 3   Discussion

Toyoaki Nishida:
I'm interested in the dynamic aspect of the semantic annotation, because it is very nice if you give a lot of semantic tags, semantic information to the world, but at the same time semantics is very, very dependent. The meaning of something could change very quickly sometimes. So at least the meaning is time dependent. Maybe you may give some tag at some time, but maybe three years later the tag can have a different meaning.

Koiti Hasida:
In the setting of semantic annotation and intelligent content, the semantics, the meaning of the content or the concepts that are there are not defined in terms of any particular ontology, but those things would depend upon the entire data out there. And also it depends upon the interaction between people and those data through things like information from people and translation. So, as the world changes, the meaning changes.

Toyoaki Nishida:
So could you tell me how you can actually make it depend on the outer world rather than on a fixed, static set of ontologies?

Koiti Hasida:
Yes. The actual meaning of the information content is exercised through the interaction between people and the content. And as this interaction changes over time due to social change, that interaction changes, reflects the change of the word. There is no technical theory about this, but it is very practical.

Michael Witbrock:
Do you see this process of annotating digital content for semantics as one that's going to be ongoing, or do you hope that this serves the purpose to provide training data so that we might learn to do automated annotation in the near future?

Koiti Hasida:
Of course, annotated data would be nice as research data for training your analysis program, contributing to artificial intelligence in coming years.

Michael Witbrock:
So do you think the prospects for being able to do automated annotations to fairly deep representations in the near future are good or poor?

Koiti Hasida:
Maybe in a hundred years, because I think it's very important to involve humans because only machines can do certain things, and only people can do certain things. And the emphasis is to combine these two very different abilities. And that is the only way to provide essentially new, nice support.

SESSION6:

Knowledge Processing

Web Scale AI Artifacts

R.V.Guha

IBM Almaden

3-5 year view

- Very large scale, very robust, but shallow AI
- Applications areas: search & data exchange

- AI Artifact: Billion concept KB
  - Broad shallow KB about particulars.
- Application: Semantic Search
  - Search based on understanding of denotation of query

- AI Artifact: Data Web
  - Large percentage of web in machine readable form
- Application: Internet as a Wet Lab
  - Programs analyzing experimental data driven of data web

Ongoing work relevant to 3-5 year view

- Research efforts in
  - Shallow parsing and information extraction from text/web with high accuracy
  - Semantic negotiation
  - Application: Use of semantics in search

- Standards
  - RDF, OWL, DAML-S, DQL, …

- Knowledge Systems
  - PharmGKB, TAP, OpenCyc

7-10 year view

- Data Web
  - Every field has its own data web
  - PhD thesis based solely on data web
  - NIH grants require data publication

- More AI in artifacts
  - Deeper understanding
  - Autonomic management
  - AI style data mining/knowledge discovery

- More for AI: great experimental test-bed for learning, deduction, analogical reasoning, …

Session 6 Panel 4   Discussion

Thomas Dietterich:
How do you envision handling changes in format over time? For instance, if we think about bioinformatics, I'm aware of some of the attempts in proteomics to integrate protein/protein interaction data which is very noisy data, from many kinds of instrumentation, and how do you combine all of that? And when a new kind of instrument becomes available, how do you add that? It seems to me that this is a part of the extensibility.

Ramanathan Guha:
In a sense, every one of these projects solves a distributive coordination problem. This is distributive coordination problem for the meaning of particular terms that are application specific, and doing it in an extensible fashion, which is why the issue of semantic negotiation and so on gets really interesting.

Eric Horvitz:
I would suggest that, for the purpose of doing science in the future, we need to think, ahead of time and deeply, about rich schemata for storing different kinds of information and to include in such schemata representations of the reliability of the information such that, even as the information ages and the tools and the sensors change and perhaps become more reliable, we can still make use of that data by folding in a consideration of its source and reliability.

Ramanathan Guha:
You need to be able to infer that. "Okay, this is from the census bureau. It's probably high quality, but it's from '97. It's not going to be accurate now", and so on.

Eric Horvitz:
Right. I'm thinking even more so about the example of the fast-evolving field of bioinformatics, where the reliability of data collected with such tools as arrays of message RNA sensors is changing dramatically and quickly. We may not want to throw out the old data, but rather combine it and leverage it in a coherent manner even when new methods become available.

Thomas Dietterich:
I think Danny Bobrow asked yesterday about the results of science are not only the data and the assertions about the data, but also arguments about the issues, the concepts, and so on. And there are formal representations of argument structure, and discussions, and decision making processes. Is there any effort to put that kind of thing in?

Ramanathan Guha:
Not right now. It's being driven by people writing programs who are finding it really difficult to get data for their programs. And so it's not meant to be the new science, it's simply a way to use a suite of tools for the existing science. But it will be good to have all this stuff.

SESSION6:
KNOWLEDGE PROCESSING

"Knowledge Systematization through
Ontology Engineering"

Riichiro Mizoguchi

ISIR, Osaka Univ.

The Future of AI Workshop
December 14 – 15, 2002

1.1 Ontology(1)

Characteristics and utility of an ontology
  KB aspects
    To help build a model by providing a specification
    Knowledge systematization
      for Knowledge sharing & better utilization
    Heavy-weight ontology
  SW aspects
    Interoperability
    Vocabulary rather than concept
    Ontology mapping/alignment/merge
    Light-weight ontology
Ontology
  as a theory of content
  as technology of content(meaning)

Need a success story of ontology engineering

1.1 Ontology(2)

Fundamental issues
    Convincing upper ontologies
    Ontology of information/representation
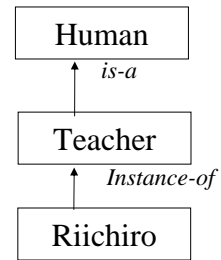   *Concept of Role*
Engineering aspects
    Ontology alignment
    Collaborative development environment
Knowledge systematization
    Articulation of the domain
    Identification of the relations among concepts
    Modeling and/or representing the knowledge in terms of Ontology

Human
is-a
Teacher
Instance-of
Riichiro

1.2 Semantic Web

VLKB
    Vast amount of data/information are out there
    How to transform it into Knowledge
    Intelligent Tag-Computing

Intelligent web services using task ontology
    Service configuration according to the task structure and
      its role assign mechanism
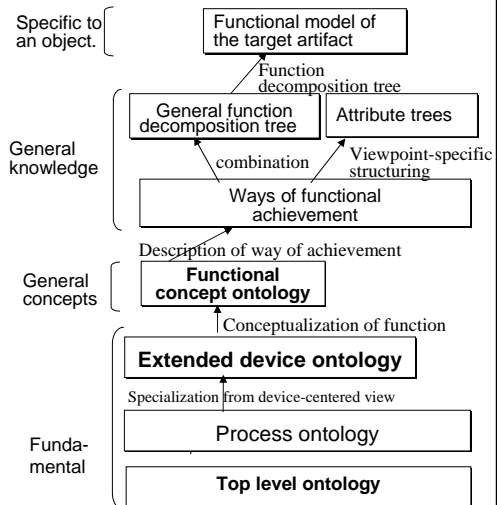
1.3 Knowledge management

KM is an issue not of a retrieval system but of a knowledge content.

Knowledge needs to be systematized for sharing

Deployment of functional ontology and a KR framework for the management at the production divisions of Sumitomo Electric Industries

The first success story of ontology engineering

One KR framework has been used for multiple purposes

A consortium is being formed

Specific to an object.
Functional model of the target artifact
Function decomposition tree

General knowledge
General function decomposition tree
Attribute trees
combination
Viewpoint-specific structuring
Ways of functional achievement

General concepts
Description of way of achievement
Functional concept ontology
Conceptualization of function

Extended device ontology
Specialization from device-centered view

Funda-mental
Process ontology
Top level ontology

---

2. Conclusion

- *Ontology engineering for knowledge systematization*

- *Future work*
  - *Systematization of nanotechnology knowledge*
  - *Systematization of Learning theory and Instructional Design theories*

Session 6 Panel 5   Discussion

Daniel Bobrow:
So is this ontology bridge nearing use for simulation of the engineering artifacts, or just for retrieval of design?

Riichiro Mizoguchi:
Not only for simulation but not only for information retrieval either. So ontology helps you how to represent the functional structure of artifact.

Daniel Bobrow:
What do you use that representation for?

Riichiro Mizoguchi:
We defined function terms in many categories of functionability. And we asked them to write it in their functional understanding of the artifact based on formalism.

Daniel Bobrow:
So is this for specification, then, and to tell whether their artifact actually meets the specification? I'm trying to understand how you would know you had the right description.

Riichiro Mizoguchi:
I just asked the engineers to write their knowledge. This is an example of a wire saw machine. We asked them to write the functional structure. But it is not only for the functional decomposition of the top level function of the machine. It contained why you have this function, how it was formed, and why you need it. So it is by conversing with them. And they checked this diagram, of course using a computer. Then we asked them to evaluate if it described the machine's functional structure well or not.

Michael Witbrock:
So, since your building a re-usable ontology, it's presumably to some extent independent of language. So one question I have is how do you connect natural language, say Japanese, to the representation in your ontology?

Riichiro Mizoguchi:
We have two levels. One is the material level, and the other one is the contextual level. We defined contextually eight terms. And we have the vocabulary to be used by these engineers.

Michael Witbrock:
So is your ontology purely term-based or do you have predicates for all those relations connected to verbs? If so, how do you take more complicated linguistic structures into your representation?

Riichiro Mizoguchi:
We don't pay much attention to the linguistic differences. This is the representation that engineers share. We don't have any natural language information from that one.

Michael Witbrock:
We're building a very large ontology. You're building a very large ontology. What do you think about the prospects of achieving a maintainable alignment between our two ontologies in order make both of them more useable?

Riichiro Mizoguchi:
That is a very good point, but at this moment we have no problems with ontology alignment because all of the ontologies are built by ourselves. If we had to share, that will be very hard for us today because we will have the share the total ontology.

Session 6   Discussion

Thomas Dietterich:
So we've been building ontologies now for maybe 25 years in AI. Is it getting easier?

Michael Witbrock:
In terms of actually adding knowledge to our ontology and booting it out, absolutely it's getting easier. At least it's getting faster. Inside Cycorp, at least, the tools that we have produced for knowledge acquisition have given us a ten-time speed up in the rate at which we can add new knowledge. And we expect that to increase another ten times in the next couple of years. For lower level representations, it's certainly becoming faster. And the reason that it's becoming faster, in part, is because we've got a lot of the upper level work done. So I think progress is definitely being made.

Stuart Russell:
I think that the guys at Cyc are doing a good job. And they have almost 20 years, 18 years of experience doing this. On the other hand, when I go out and look at the places where they have repositories of supposed semantic web ontologies for various application areas I don't see particularly high quality ontologies there. I see a lot of the same kinds of errors that McDermott complained about in his "AI Meets Natural Stupidity" paper. I see confusion between instances and categories, and so on. It's my job to try to educate people on how to do these things better, but other than saying, "Well, you just have to practice and you have to do it for 18 years," there's not a lot of very precise guidance we can give. And I spent a long time looking for papers where someone could, for example, give you any kind of precise guidance on when something should be an individual entity and when it should be a concept, and didn't find that kind of information.

Riichiro Mizoguchi:
Nicola Guarino has good heuristics that he calls "odd clean". If you follow his algorithm, then you could have some beautiful and consistent top level problem.

Michael Witbrock:
And there is another approach to that. The approach that we have taken when dealing with normal people is that you just can't tell people to reliably distinguish between individuals, but you can teach the computer program to do it. So that's what we try and do. We try and just let people make these mistakes and then work out ways of fixing them, identifying when these mistakes have consequences, identifying what those consequences are, and then retrospectively fixing the ontology.

Stuart Russell:
That's hopeful, but I mean if people go right to these semantic web ontologies, the components of a semantic web, they're the world's leading experts on ontology and so on.

Michael Witbrock:
But even internally we can't do it reliably. How do you tell that a disease is a collection and not an individual? There is no good philosophical basis for making that choice sometimes. And sometimes that's just what you do.

Daniel Bobrow:
Is there a functional basis? That's why I keep asking, "What are you doing with this?" What we saw here was being used in communication about artifacts. And so being a communication medium you can count on the human interpretation and it probably makes less difference. What functional specification do you have of whether it's good to be an individual, or a concept for things that you want to have that distinction make a difference?

Michael Witbrock:
Well, for example, if you're going to have individual instances of them it should better not be an individual. So there are simple functional criteria.

Daniel Bobrow:
But that's a philosophical thing, which is what Stewart was asking. I'm asking a more pragmatic thing, which is "Why would I want to have an individual of a disease, or when would I want to have that?" What are you doing in terms of reasoning that you would like to be able to sometimes create individuals and sometimes not? Can you say some things that are functionally happing with this that may give you guidance?

Stuart Russell:
I think a good start would be in a casebook. Here is a set of different applications, and here are a set of attempts to build an ontology that covers all of these applications, and here is why such and such an approach doesn't work, and here is why it does. So if you can't derive general principles then at least we should come up with examples.

SESSION7: "SYNTHESIS, SUMMARIES, RESPONSES and other TOPICS"

Koichi Hori

University of Tokyo

The Future of AI Workshop

December 14 – 15, 2002

Super Intelligent AI ?

or

AI to make humans super intelligent ?

Question of personal preference?

Question of social accountability?

Question of social accountability

Biologists are good at explaining (or exaggerating) the social value of their research.

(It is easy. They just should say that their research will lead to new medicine.)

How about AI?

We should claim that AI can be medicine to cure the disease of human society.

Feigenbaum test vs. Hori test

The human society which lived with AI will be more creative and more robust than that which lived without the AI system.
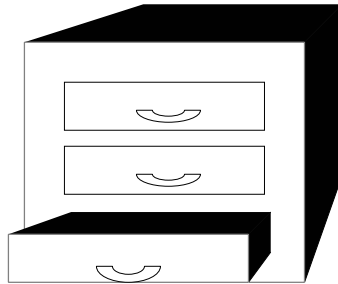
2. Conclusion

`Super Intelligent AI' and `AI to make humans super intelligent' may be two aspects of the same thing.

Enhancement of Knowledge Creation Communication Cycle

Koichi Hori

(University of Tokyo)

Myth of Knowledge Management

- Knowledge is some 'thing' that can be stored, retrieved, and reused.

Many complain that:

- We cannot capture knowledge,
- We cannot maintain knowledge,
- The stored knowledge is useless,
- ......

The reality is
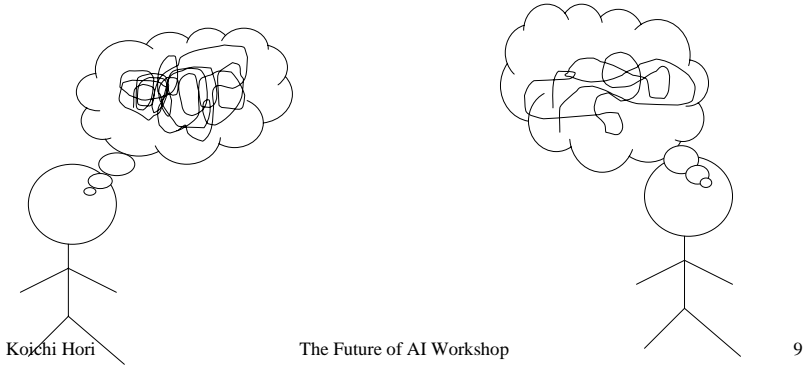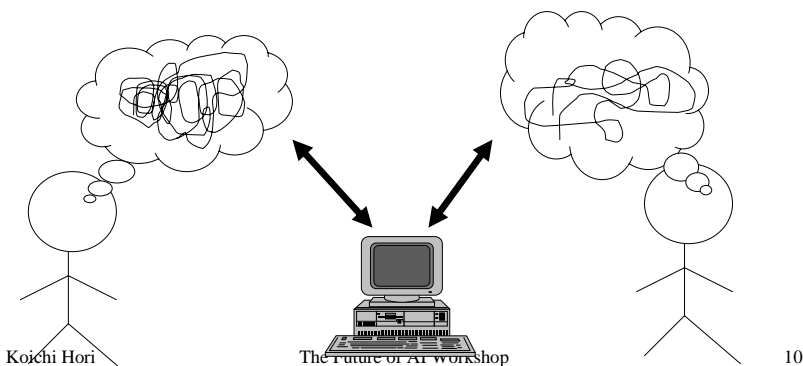
- Knowledge is not a thing, but
- knowledge emerges dynamically from the interaction among nebulous mental worlds.

What computers can do is

- to enhance the interaction among the nebulous mental worlds, and
- to stimulate the nebulous worlds.

To stimulate the nebulous world,

- Computers can collect the nebulous source.
- Computers can store the nebulous source.
- Computers can re-crystallize the nebula.

Using the technologies of

communication monitoring, ubiquitous computing, data mining, natural language processing, semantic annotation, visualization, interaction, and so on.

Knowledge Nebula Crystallizer

Supports dynamic knowledge emergence through re-organizing stored nebulous knowledge source depending on new contexts.

Session 7 Panel 1   Discussion

Ron Brachman:
We're trying to do in one place something important, and that this audience would appreciate, but in fact we still have many, many arguments to make with, for example, the United States Congress. And there are people who not only, just as you suggest in Japan, not only don't appreciate AI, but don't even appreciate the importance of what you might call information and computing technologies at all. And we have the very same problem, that the biologists in the United States are much more effective at arguing for very large amounts of money than anyone else. So there may be a little gleam of hope, but it's not as if we're doing things totally different in the U.S. government.
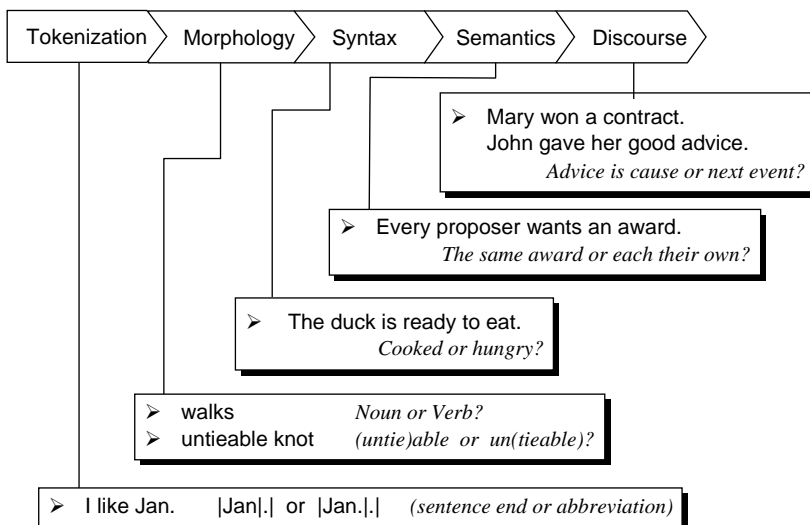
Session 7: Synthesis, gaps, challenges

A missing issue:
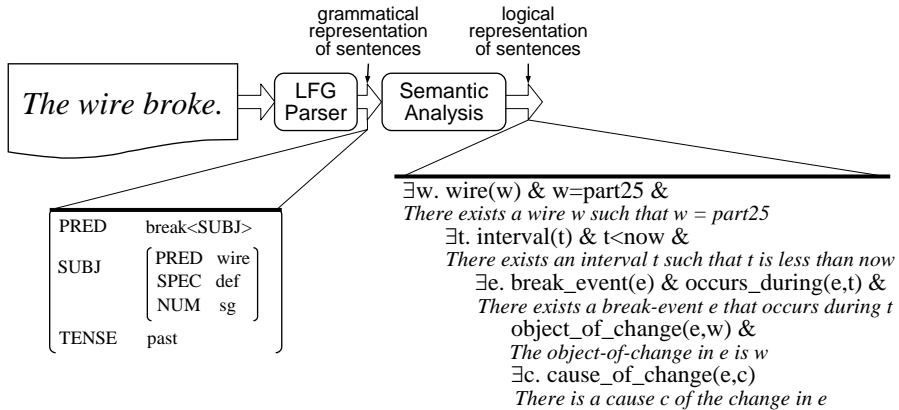**Problem of language understanding**

Some recurrent themes

Daniel G. Bobrow
Palo Alto Research Center
December 2002

parc

Ambiguity is Pervasive

Tokenization 〉 Morphology 〉 Syntax 〉 Semantics 〉 Discourse

➢ Mary won a contract.
   John gave her good advice.
      Advice is cause or next event?

➢ Every proposer wants an award.
      The same award or each their own?

➢ The duck is ready to eat.
      Cooked or hungry?

➢ walks   *Noun or Verb?*
➢ untieable knot   *(untie)able or un(tieable)?*

➢ I like Jan. |Jan|.| or |Jan.|.| *(sentence end or abbreviation)*

parc

Semantic Representations

grammatical representation of sentences | logical representation of sentences

The wire broke. → LFG Parser → Semantic Analysis

```
PRED     break<SUBJ>
SUBJ     PRED    wire
         SPEC    def
         NUM     sg
TENSE    past
```

$\exists$w. wire(w) & w=part25 &
There exists a wire w such that w = part25
  $\exists$t. interval(t) & t<now &
  There exists an interval t such that t is less than now
    $\exists$e. break_event(e) & occurs_during(e,t) &
    There exists a break-event e that occurs during t
      object_of_change(e,w) &
      The object-of-change in e is w
      $\exists$c. cause_of_change(e,c)
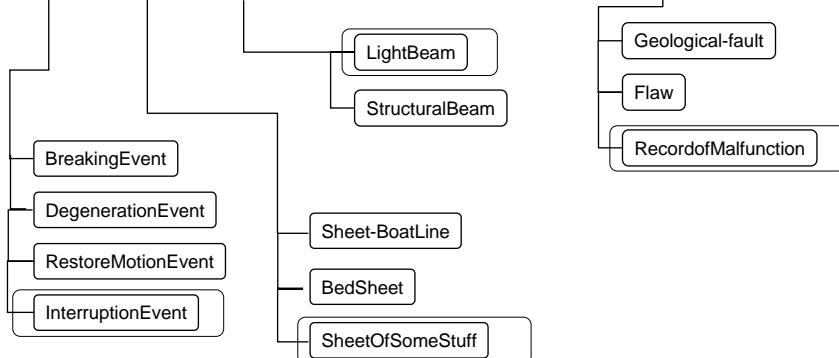   *There is a cause c of the change in e*

- LFG F-structure gives basic predicate-argument structure, but lacks:
  - Standard logical machinery *(variables, connectives, etc)*
  - Implicit arguments *(events, causes)*
  - Contextual dependencies *(the wire = part25)*

parc

Mapping into Concepts

The sheet breaks the beam too late, causing a fault.

`(breaks sheet beam) (too late) (cause x fault)`

LightBeam
StructuralBeam

Geological-fault
Flaw
RecordofMalfunction

BreakingEvent
DegenerationEvent
RestoreMotionEvent
InterruptionEvent

Sheet-BoatLine
BedSheet
SheetOfSomeStuff

Evidential reasoner chooses likely mutually compatible choices
Constraints:
  Ontological: *InterruptionEvent* compatible with *LightBeam*
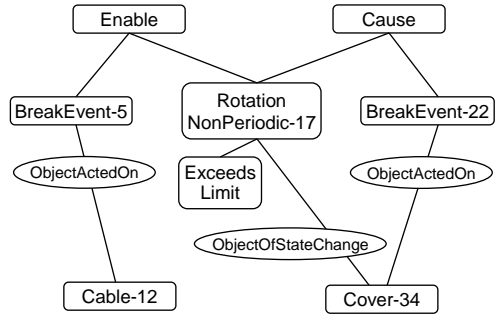 **Common sense: *SheetOfSomeStuff* unlikely cause of *BreakingEvent***
  Domain specific: *fault* almost always a *RecordofMalfunction*

parc

173

Canonical Representation of Content

The cable is breaking, allowing the cover to pivot too far, breaking the cover.

The cable fails, which lets the cover open too wide, causing it to crack.

Enable

Cause

BreakEvent-5

Rotation NonPeriodic-17

BreakEvent-22

ObjectActedOn

Exceeds Limit

ObjectActedOn

ObjectOfStateChange

Cable-12

Cover-34

parc

A challenge: Knowledge Fusion
Getting the document you wish had been written

Tip 27009

Problem: Compiler ~~ing intermittently~~

Solution: ~~the fault cannot be~~ ed do the following ~~first:~~ Enter D~~30, o~~ the finish~~op cover~~ che~~e interlock~~
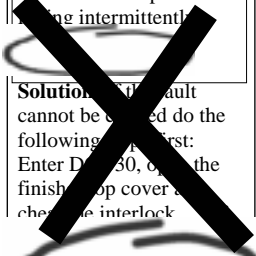
Tip 27011

Problem: FC 12-312

Cause: . . .The tin plating on the sensor pins oxidizes and creates a poor connection.

Solution: ..."slide" the connectors on and off the sensor pins. This will scrape the pins of the sensors and provide a better contact. Long Term: Sensors with gold plated

Tip 27018

Problem: Intermittent faults in compiler eject/compile positions

Cause: Bi-metallic ~~corrosion~~

Solution: Use Kit #600K455. It contains sensors (4) with gold plated pins to replace the Eject sensor.. . .now in parts supply.

New Tip 27011/018
 Problem
 Cause
 Solution

parc

174
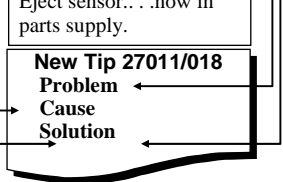
Improving the Feigenbaum Test

- Principle: Divide and conquer – focus on expertise
  - Test a real expert against an AI expert
- The hidden assumption:
  - Language is simply built on knowledge
  - Social interactions will not affect judgement

- Improving the test
  - Use two a non-native speaker expert ??
  - Pair a graduate student with the expert (both human and AI)

parc

Recurrent themes of the workshop

- Integration is important
  vs. we can't have everything at once
- Representation is the key
  vs. intelligence is about action, and interaction
- Our goal should be a complete AI
  vs. intelligent assistants and interaction
- Focus on the internal resources of an AI program
  vs. the use of external resources (social, artifacts)

parc

Session 7 Panel 2   Discussion

Manuela Veloso:

I think that challenge that you showed about coming up with a new tip, it's really compelling. So do you believe that this is feasible, even with the methods we have now, the parsers, the semantics, the Cyc's, the semantic web, the whole thing?

Daniel Bobrow:

I believe that within the next 10 years, in some domains, we will be able to do that. I think the kind of technology that we are working out on a five-year program will get us to do some of these things, because what it takes is actually understanding multiple levels of structure in what you're seeing, and the fact this was a work-around, and that was the final solution because it mentions the real part that you should replace it with. There are ways of doing this. I don't think this is a generic kind of thing, but then again, I'm terrible at writing summaries of multiple papers, too. But under certain circumstances I can do that sort of thing. And I think you should try to do much better than they do.

Session 7:

synthesis, summaries, and other topics

Focus

- assemble the messages towards the outer world

  - The last opportunity to discuss; but let us move our eyes towards the outside (AI people not here + the outer world)

  - Identify missing issues, and future challenges in particular.

Collective Intelligence or Social Intelligence

- the group's ability to deal with complexity, by capturing, sharing, extracting meaning from signals

- The individual's ability to interact with other agents in the society

Collective Intelligence or Social Intelligence

Conversational systems:
 that can communicate with people
 or that mediate conversation with people

DEMO
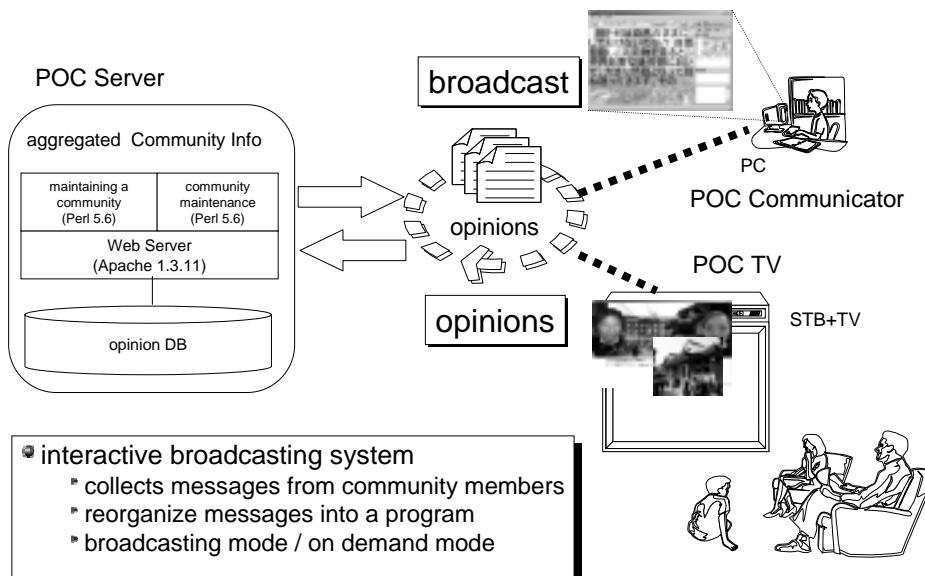virtualized egos:
conversational agents that
talks on behalf of me

demo (NL interaction)

[Kubota 2002-09-13]

FTTH Trial (with KDDI)
-- Public Opinion Channel for Local Community

POC Server

aggregated Community Info

maintaining a community (Perl 5.6)	community maintenance (Perl 5.6)

Web Server (Apache 1.3.11)

opinion DB

broadcast

opinions

PC
POC Communicator

POC TV

opinions

STB+TV

- interactive broadcasting system
  - collects messages from community members
  - reorganize messages into a program
  - broadcasting mode / on demand mode

## Human-artifact Communication -- mobile chair agent

Issues in Social Intelligence Design

- methods / tools of establishing the common ground
  - ◆ sharing/circulating stories
  - ◆ sharing sense of awareness
- agent mediated communication
  - ◆ embodied conversational agents
  - ◆ socially intelligent agents
- collaboration design
  - ◆ bricks/bits/interaction
- public discourse and e-democracy
  - ◆ conflict resolution and negotiation assistance
  - ◆ survey, Delphi and mediation
  - ◆ visualization
  - ◆ open discussion, public opinion process and decision making
- Evaluation and measurement
  - ◆ log analysis
  - ◆ social network analysis
  - ◆ social intelligent quantity

Some social questions

- Artifacts that can communicate with people with multiple communication means
  - ◆ Verbal communication / nonverbal communication
- Agent mediated communication
  - ◆ Can it help?
  - ◆ Can it bring about intellectual interaction
  - ◆ Can it help increase the creativity of the entire community
- Information summarization
  - ◆ may have critical social implication (e.g., likelihood of earthquake, economical issues (which may affect stock market), community decision making (distribution of opinions), …)

AI as developing an automobile

Automobile

steering

sheet

The Engine

tires
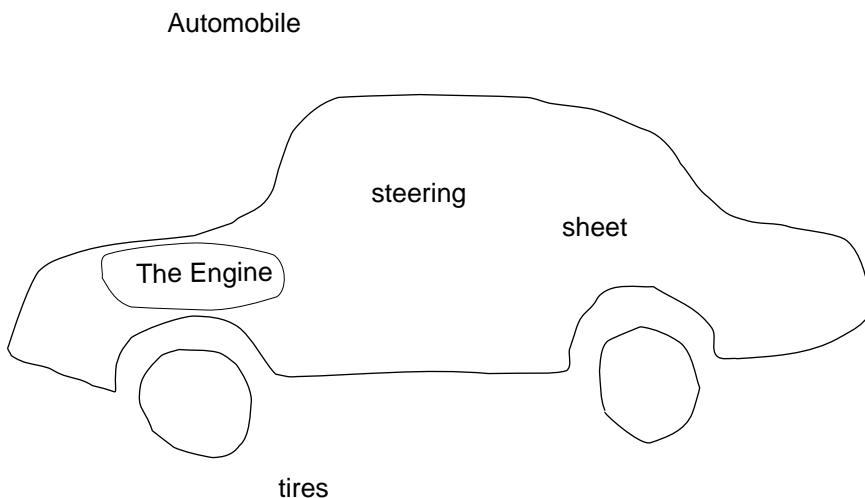
My personal experience

- Serious AI-related research outside the AI community.

- We should look at non-computational aspects, as well as computational aspects of intelligence.

FAIW agenda

- keynote speech: Feigenbaum
- session 1: foundation of AI chaired by Sakurai
- Arimura, Russel, Ueda, and Sakurai
- session 2: discovery chair by Motoda
- Suzuki, Miyano, Dietrich, and Motoda
- session 3: HCI chaired by Horvitz
- Sumi, Nakakoji, Ishida, and Horvitz
- session 4: AI systems chaired by Feigenbaum
- Brachman, Washio, Sasaki, Ohsuga, and Feigenbaum
- session 5: human-level intelligence chaired by Miyake
- Ohsawa, Suwa, Veloso, and Miyake
- session 6: knowledge processing
- Nagao, Witbrock, Hasida, Guha, and Mizoguchi
- session 7: synthesis, summaries, responses, and other topics
- Hori, Nishida, Bobrow, Cohen

Some focal points

- **Explicit**
  - ◆ modified Turing test or Feigenbaum test?
  - ◆ mathematical AI or AI as system integration or ...?
  - ◆ from symbolic (or QA) AI to situated AI?
  - ◆ artificial situated cognition possible?
- **Tacit**
  - ◆ AI vs HCI

Session 7: Synthesis, gaps, challenges:
Semantic Autonomy

Paul Cohen
University of Massachusetts

Outline

- What is semantic autonomy, why do we want it
- How do we get it – a sketch and prototypes
- How do we get it – research challenges. Here I will identify themes from the workshop.
- One challenge problem and one grand challenge problem

- Today we engineer representations so that syntactic processes will manipulate symbols that mean something to us, to produce symbols that mean something to us

- It's irrelevant whether the symbols mean anything to the machine

Semantic Autonomy: Meanings
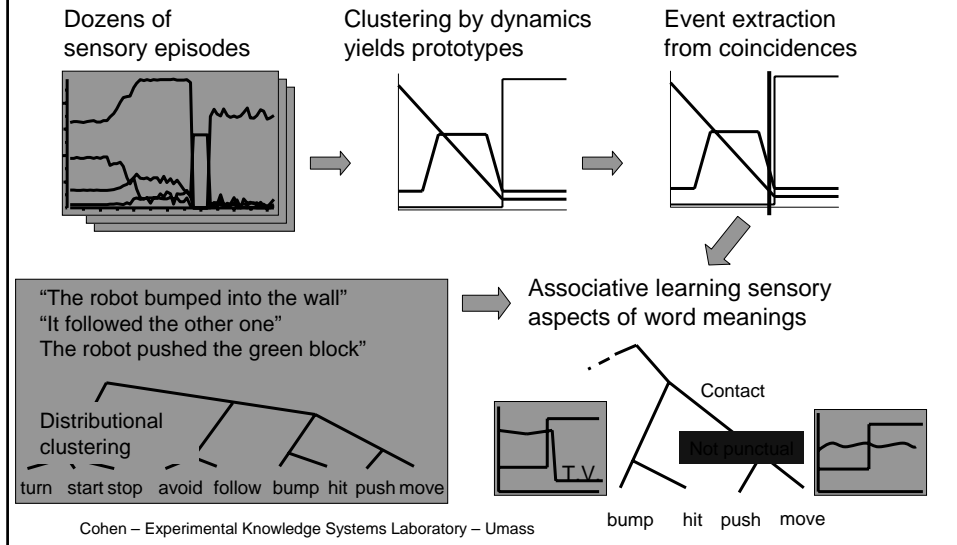of representations are learned
by and for the machine

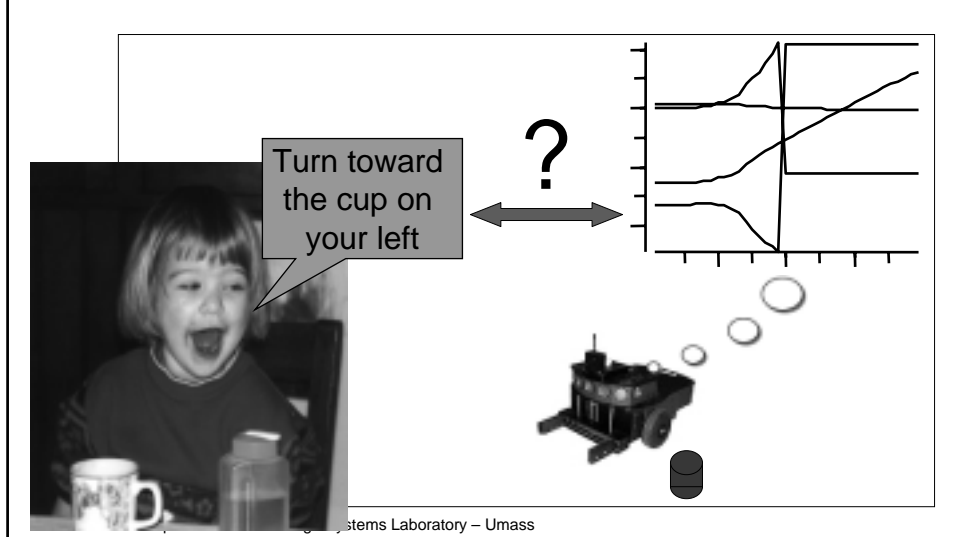Why we need semantic autonomy

- People have it
- Semantic babysitting – the business of checking that symbols in the machine mean what we want them to mean, and debugging otherwise – is expensive

Semantic Autonomy: Unsupervised learning of meaningful representations from sensory data (grounding)

Dozens of sensory episodes

Clustering by dynamics yields prototypes

Event extraction from coincidences

Associative learning sensory aspects of word meanings

"The robot bumped into the wall"
"It followed the other one"
The robot pushed the green block"

Distributional clustering

turn start stop   avoid follow bump hit push move

Contact

Not punctual

T V

bump   hit   push   move

The trouble with sensory prototypes – they don't denote objects and they aren't compositional

Turn toward the cup on your left

?

186

We can cluster sequences of propositions by dynamics, too.

Turn toward the cup on your left

(cup A)
(red A)
(left-of R A)

Cohen – Experimental Knowledge Systems Laboratory – Umass
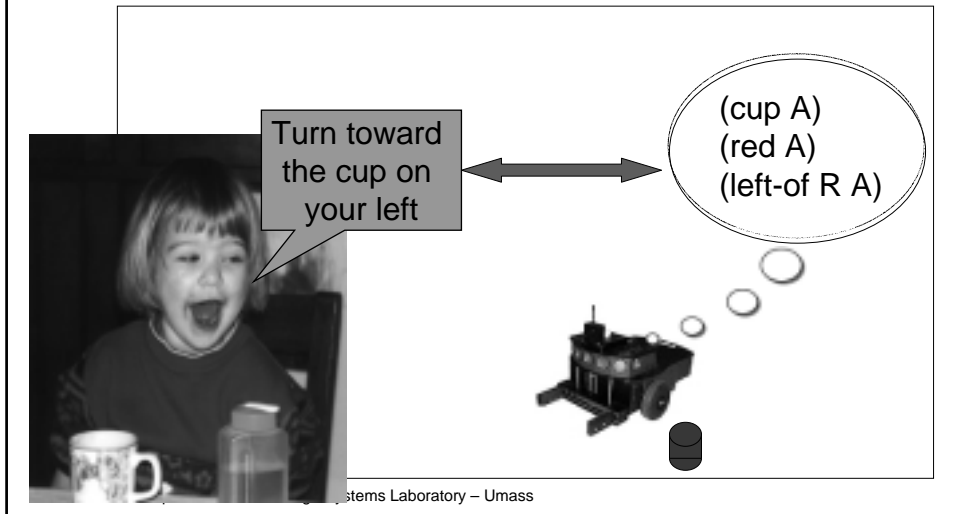
Challenge Problem: The enrichment or deepening of knowledge through experience (thanks Michael Witbrock)

Sensory prototypes from experience

forward movement

stopped (with wheel slippage)

Translational Velocity

Forward Sonar
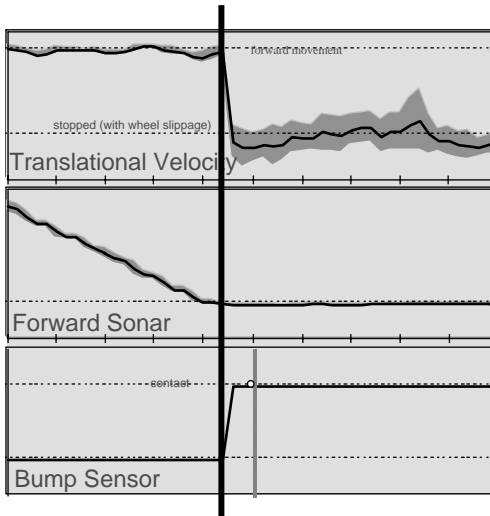
contact

Bump Sensor
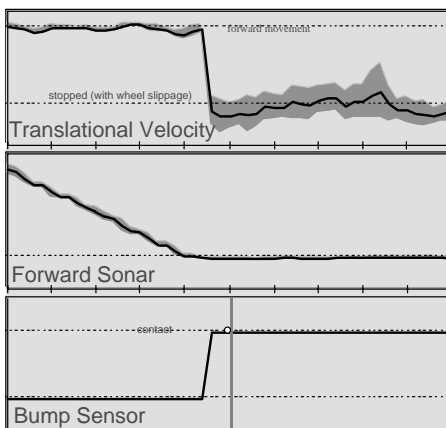
Cyc knows a lot about distance, including:

- A salient fact about *touches directly* is that iff some X and some Y touch each other or share a part then the shortest distance between X and Y is zero feet.

Can these representations enrich each other and the sum of the knowledge?

Does grounding help?

Cohen – Experimental Knowledge Systems Laboratory – Umass

Extending the knowledge
From prototypes to axioms and axioms to prototypes

forward movement

stopped (with wheel slippage)

Translational Velocity

Forward Sonar

contact

Bump Sensor

Suppose the concept *collision* exists in the Cyc knowledge base. It can be found by matching the following description to its axioms:

- Before this event, translational velocity is positive, afterwards it is roughly zero
- Before this event, a distance-measuring device returns a positive number, afterwards it returns zero
- Before this event the bump sensor is low, afterwards it is high

Then any assertion about collision can be conjectured to hold for these prototypes, e.g., "bump sensor" measures (in part) change of state from not touching to touching

Cohen – Experimental Knowledge Systems Laboratory – Umass

Active learning and meaning autonomy

forward movement

stopped (with wheel slippage)

Translational Velocity

Forward Sonar

contact

Bump Sensor
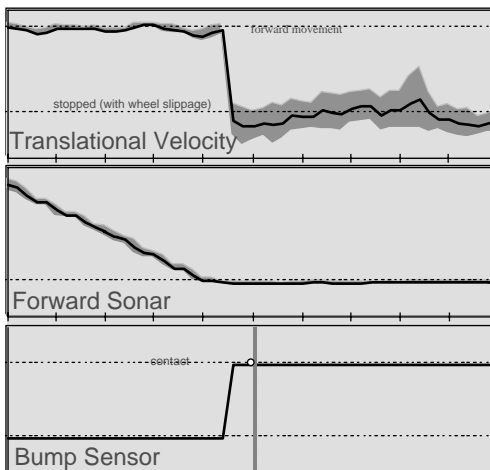
The prototype can also generate conjectures about *collision*, such as, "every collision is preceded by a reduction of distance to zero"

This is a conjecture about the meaning of the word "collision", generated by the machine! It's trying to figure out what a word means!

Cohen – Experimental Knowledge Systems Laboratory – Umass

Themes - I

- Semantic autonomy: Learning meaning by and for machines.
- Semantic autonomy does *not* mean the machine's meanings are independent of yours. How can we have semantic autonomy *and* shared meaning?
    - The problem of shared meaning, mind-reading, and HCI, CSCW; Ishida, Nakakoji, Horvitz, Brachman, Washio
    - Getting the right representational primitives (e.g. deictic markers) (Dietterich, Motoda) so robot representations and meanings can be learned to correspond to ours, but aren't designed to
    - Robots live in the same world and experience it in *similar* ways. Enormous structure in multivariate time series. Sakurai, Veloso, Cohen. Conjecture: basic ontology and knowledge is *inevitable*.

Themes - II

- Integrated perception, reasoning and action: Some meanings are about contingencies between perceived state and action outcomes; learned associatively; (Hasida, Cohen, Veloso). More likely to have shared meaning if we have shared experience.
- Deepening, elaboration, enrichment of knowledge Witbrock, Feigenbaum. Human knowledge is much deeper than machine knowledge
- Social environment, collaborative tasks, enable meanings (particularly word meanings) to be learned associatively; "communication is the process of establishing a meaning" Nishida and Hasida, Sumi, Miyaki, Brachman,
- Methodological: Focus and integration. Brachman, Feigenbaum; Challenge problems and Turing Tests

Feigenbaum tests and Turing tests

- The Turing test was a proxy: No-one could pass the test and not be able to pass indefinitely many other tests. By design, not true of "partial intelligence" (Feigenbaum) tests. Work on complete AI.
- But failure on any "I know it when I see it" test (e.g., the Turing test) is uninformative
- We need *diagnostic* tasks. Failure must inform progress.
- Hori test doesn't do it but is the first interesting variation I've seen on Turing's proposal

Grand Challenge: Robots that learn human language

- The twin problems of semantic autonomy and shared meaning
- Perceptual learning (sensor-to-symbol; what are the innate representational primitives; what is the syntax and semantics of constructions; how does the conceptual system – ontology and axioms – on which language is layered develop?)
- Grounding semantics (particularly verbs and prepositions) in interaction with the environment
- Social learning (you know the meanings of words, I don't; how can I learn them by listening and watching and acting?)
- Active learning (I think I know what a word means or how to construct a phrase, but I'd like to deepen or test my knowledge)
- Compositional semantics, LOT etc.
- What is the role of language in modifying the conceptual system
- Foundations: what are mental states, how can they have meaning, does meaning influence behavior, how do lexical semantics become associated with these states, is associative learning sufficient?

Concepts from sensory data
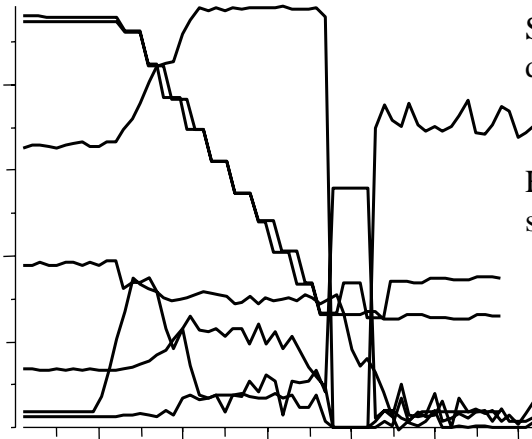
Start with unlabelled data from robot sensors.

Here is one instance of six seconds duration.

Concepts from sensory data
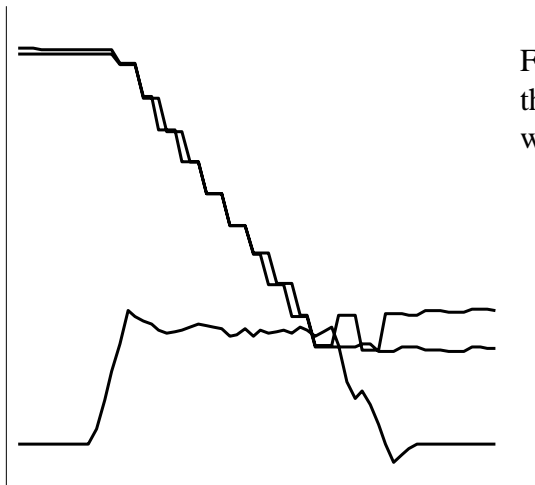
Find subsets of sensors that change in similar ways across experiences.

Clustering by Dynamics: Similarity

- Clustering methods are unsupervised, and group together elements based on their similarity

High intragroup
similarity

Low intergroup
similarity
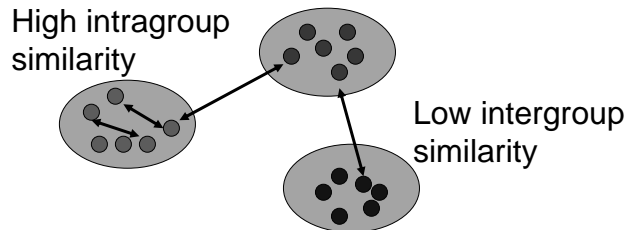
- How to judge the similarity of multivariate time series ?

Multivariate series in each cluster are averaged to produce a *prototype* for the cluster

forward movement

stopped (with wheel slippage)

Translational Velocity

Forward Sonar

contact

Bump Sensor
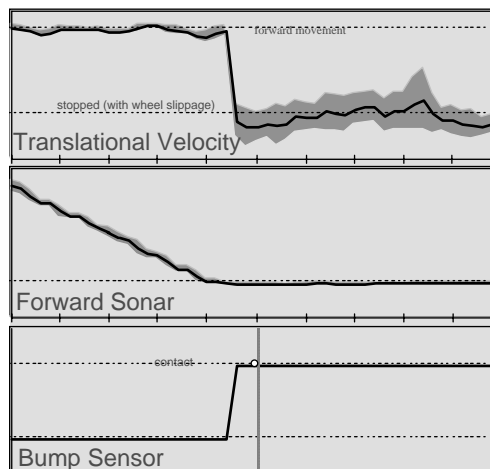
Using sensor prototypes as semantics for distributional clustering of words in descriptions of robot activities

Pattern i distinctive for punctual because t.v. drops

Events (occurrences of words or word classes)

Contact

punctual | Not punctual

bump | hit | push | move

Pattern j

Cohen – Experimental Knowledge Systems Laboratory – Umass
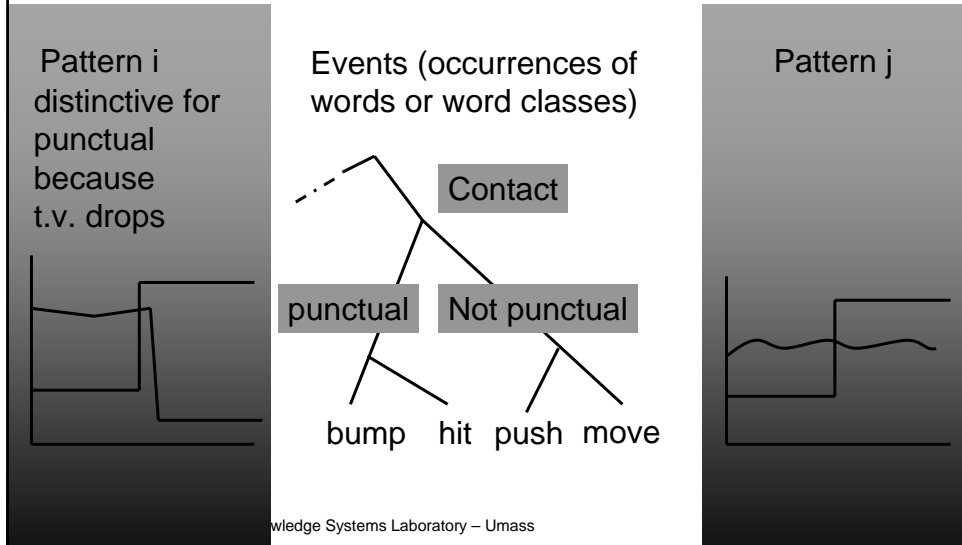
Challenge Problem: Cycbots

- Cyc is not integrated with a perceptual system or an action system
- Grounding should help
- Integrate Cyc with a robot to enrich and deepen its knowledge

indicates

imminent crash

sonar-0 < 20 ——— causes ———→ stop!

structuring cause

explains

triggering cause
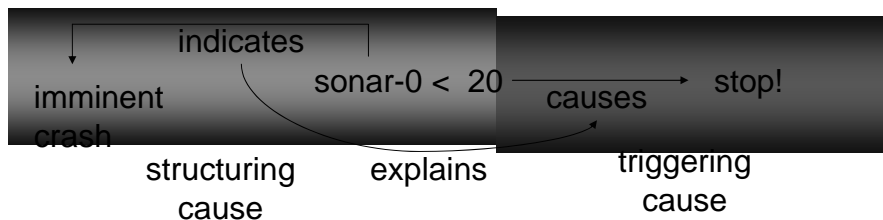
The meanings (or contents) of states

indicates

F: temperature   C: bimetallic strip   causes →   M: furnace goes on

explains
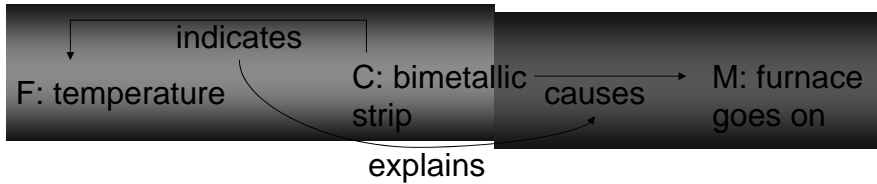
- The meaning of C is that it has the function of indicating F.
- The meaning of C is *not* that it has the function of causing M; in fact, C is only part of the cause of M (there's also the desired temperature setting)
- "C acquires its semantics, a genuine meaning, at the very moment when … its natural meaning (the fact that it indicates F) acquires an explanatory relevance." Dretske, Explaining Behavior, p.84

194

Session 7 Panel 4   Presentation

Paul Cohen:
Well, I'm Paul Cohen. I'm from the University of Massachusetts. And I have the honor of being the last speaker here at this workshop. And it is an honor because it's been a remarkable workshop. And I've had the opportunity to listen to a great many ideas. One of the things I'll try and do this afternoon is synthesize some of those ideas and try and identify some themes.

My talk is called "Semantic Autonomy." And what I'll start by doing is describe what I mean by semantic autonomy and why we would want it. And then I'll give some examples from my own work about how you might get it. But you'll see that my own work is very preliminary. And so in identifying the research program to provide agents with more semantic autonomy I'll have the opportunity to reflect on many of the good ideas I've heard here over the last couple of days. And I'll conclude by issuing a challenge problem. Actually, I'm going to issue one challenge problem in the middle and one grand challenge problem at the end.

Well, today in artificial intelligence we engineer representations in such a way that syntactic processes can manipulate symbols that mean something to us. And when they're done manipulating those symbols what comes out is something that means something to us. Nobody asks, because it isn't really considered relevant, whether those symbols mean anything to the machine.

The idea of semantic autonomy, on the other hand, is that meanings of representations are learned by the machine for the machine. And I have a couple of photographs here. Those of you who know me know that I rarely pass up an opportunity to show people how talented, and beautiful, and intelligent, and just generally wonderful my daughter is. These photographs are very meaningful to me. They're probably somewhat less meaningful to you. But there is still shared meaning there. You can tell, for example, that there is a kid here shooting a bow and arrow. You can tell that she's hit the target. You might be able to identify a wild pig.

You wouldn't know that the wild pig is important because it's about the only meat that my daughter will eat. So there's some intersection between what we all know when we encounter a symbol. There's some intersection between the semantics of symbols that we have in common and that which is personal.

But the point I really want to make about these pictures has to do with the look of sheer, unadulterated joy on this little girl's face when her arrow pierces the heart of that wild pig. It's a look of great victory, and it means something to her that she can shoot well enough to kill a paper wild pig. And she's learned that herself. I never told her what it meant. I've not told her about the thrill of victory and the agony of defeat. This is something she has learned for herself.

So I'm after agents to have semantic autonomy, where semantic autonomy means that the meanings of symbols are learned by and for machines. And the reasons I want it are that people have it and that semantic babysitting, the process of checking the symbols in a machine, mean what we want them to mean. And debugging the machine otherwise is enormously expensive.

195

So here's a sketch of one way that we've been able to achieve semantic autonomy for a robot. And I want to stress, first, that it's a sketch, and second that it is one of several ways that we've been able to accomplish this. I don't really want to stress particular algorithms, nor are the algorithms that I'm going to describe the most advanced that we have. But the sequence will make my point.

It is in fact possible to start with a robot and collect dozens of time series, multi-variant time series from the robot's sensors, and then using unsupervised clustering algorithms come up with prototype source of the average members, average representations, of sequences of sensory signals. And in this case you see the average member has this sensor going down to here and then flattening out. And this sensor ramping up and going flat for a while, and then ramping down and flattening out. And this sensor remaining low and then going high. And it's relatively easy to get patterns like this if you start with raw sensor data.

Indeed, because you are doing clustering what you will often find is several prototypes, several average members, several sort of central members. And it turns out that those correspond to episodes in the life of the robot. It's also relatively easy to look for places where things change simultaneously and therefore extract events in the life of the robot.

Now at the same time, at least if you're a human being, at the same time that you're doing things people are talking. People are saying things. There's a language stream going on, in addition to your action stream. And there are a number of kinds of unsupervised learning you can do with a language stream. One of the simplest is called distributional clustering. And in distributional clustering you build a hierarchy of words using the notion that words that occur in similar contexts are likely to be similar words. They're likely to have similar meanings.

Linguists often build hierarchies like this. But linguists always label the internal nodes themselves. They come along and they say, "that's a direct object," or "this is past participle," or "this is verb," for instance. What we were interested in is whether we could learn the meanings of these internal words by associating them with these sensory patterns that the robots learn. And indeed, we can.

It turns out, in an experiment we ran in which we asked graduate students to describe what the robot is doing in unrestricted tests. And we built sentences like that, and the distributional clustering like that. It turned out it was relatively easy to associate patterns in such a way that they could discriminate one sub-tree of this hierarchy from another.

And so here's an example. It turns out that "bump" and "hit" are associated with patterns in which the translational velocity drops to zero, whereas the words "push" and "move" are associated with patterns in which the translational velocity fluctuates but doesn't really drop. And I would claim that this is a very primitive semantics for these words. It's a semantics that's been learned in an entirely unsupervised way using nothing more than simple clustering techniques.

So I think it is possible for agents to learn the meanings of words by trying to relate those words or sentences to what they're doing in the world. It turns out that that particular kind of representation, the sort of squiggly line representation or sensory prototype representation, won't get you very far. One of the reasons for that is that they don't denote objects. And so if you are trying to

196

understand the phrase "turn toward the cup on your left," it's hard to know what you hear denotes the cup on your left. So they're not prepositional, they're not denoting, they aren't compositional. But, nevertheless, we have been able to learn the meanings of words in a non-supervised way relating to sensory data. Lately we've been working with a real propositional time series and we can perform pretty much the same trick there.

What I'd like to do now is shift gears and talk about how we might go further. And I'd like to thank Michael Witbrock very much for helping develop this example with me last night. This is a representation that has been learned in a non-supervised manner by the robot. Leave this stuff out. Don't pay attention to this stuff for just a moment. This is what the robot has learned. There is a prototype, or something that happens frequently in the life of the robot. And I think you'll be able to see what's going on.

Translational velocity is high, and then it drops abruptly. The forward sonar goes to zero and stays there. And the bump sensor goes high. And these three things happen simultaneously. Now, if I tell you what these sensors are, if I give you an interpretation of these sensors, you know exactly what's happened, right? The robot has just crashed into the wall.

The robot, of course, doesn't have an interpretation of those sensors, doesn't know anything about walls. And so the question is, could we use propositional information like Cyc's knowledge—and Michael found this yesterday in the Cyc knowledge base—the salient fact about touches directly, is that if some x and some y touch each other, or share a part, then the shortest distance between x and y is zero feet. Could you use a propositional representation like that in concert with a representation like this to kind of bootstrap the amount that you know about the world, to accelerate learning.

So the challenge problem I'm going call the enrichment or deepening problem. And by enrichment I mean that we're going to enrich what we know about the world as is inherent in this kind of representation and this representation by playing them off against each other. Another way of asking that question is does grounding actually help Cyc?

So the first step would be to identify the simultaneity of these things. Again, we have algorithms that can do that. Suppose the concept of collision already exists in the Cyc knowledge base, but we don't have a notion of collision here. That is, we don't have a sensory notion of collision. Yet we do have this prototype. We'd like to bring them into association. We'd like to say they're the same idea.

Well, one way to do that is to extract from this a declarative representation, an event-based representation like, for example, before this event translational velocity is positive, afterwards it's roughly zero, and so on. And then suddenly search the Cyc knowledge base for things for which this description is true. And then all of the assertions or axioms that hold of those things for which this description is true can be hypothesized to be true of this description also.

So that's one way that we could enrich. That is, we could enrich by taking something that we know declaratively and enhancing a representation like that with those propositions. Another way we could do it is to go in the other direction. A picture like this can also generate conjectures about a declarative notion of collision. For example, every collision is proceeded by a reduction

197

of distance to zero. It didn't say that in Cyc, but it does say it here. So if you have this kind of representation you can bring them to correspondence. You can start to ask, "Is this something that's true?" I mean, is this something that's true in the world?

Well, what I'd like to point out about that is that it's conjecture about the meaning of the word "collision" which is generated by the machine. The machine is asking what a word means. The machine is trying to elaborate the meaning of the word. That's a completely different ball game than Cyc engineers spending all of their time saying *a priori* what words mean. And that's why I think this idea of semantic autonomy, or machines trying to acquire meanings by and for themselves is actually pretty helpful.

But—so I've given a couple of examples, one that runs in my lab, one that is hypothetical that Michael helped me with. But now let's talk about how we go further. If you're after machines that learn meanings by and for themselves you have the problem of shared meanings. It's a problem that has interested philosophers for a long time. The problem really is that we want a machine to acquire meanings for symbols that are personal in the sense that they're their own meanings. But we also want shared meaning. We don't want semantic chaos.

And the problem of shared meaning has come up several times in this workshop. At lot of people have been concerned about what we might call the mind reading problem in HCI. How is it that a machine knows what I have in mind? How is that we can actually mean the same thing by symbols? We've seen the same issue come up in computer supported cooperative work. And while these names are supposed to be indicative of talks in which these issues came up, I don't want anybody to feel like I've left them out. If your name isn't on any of these lists please don't take umbrage.

So the problem of shared meaning is an important one that will become more important if machines are given responsibility for figuring out the meanings of things. Another thing that you can do to try and insure that we're on the same page about the meaning of something is get the right representational primitives. We use deictic markers. It doesn't really matter.

The point is—Tom Dietterich pointed it out, and Motoda-san pointed out that we engineer representations currently. And we will probably always have to engineer representations at some level. But what we're really after is a well-engineered set of representational primitives such that the machine can take it the rest of the way and when the machine takes it the rest of the way it doesn't diverge dramatically from what we think. And that's a research challenge.

Another way that we can hope for shared meaning with our robots is that robots live in the same world as we do and experience it in similar ways. Now, Manuela showed a marvelous movie of what the world actually looks like to a robot. And so you may think that I'm crazy in saying that they live in the same world and experience it in similar ways. But, you know, there are surfaces. There are gradual changes over time. There are abrupt changes. There's size constancy. That ball that you saw getting bigger and smaller, it turns out that the human perceptual system will always perceive as the same size. We really do live in roughly the same world and experience it in roughly the same way. I'm glad to see Manuela is nodding.

The other things is that the world itself has, especially if you take a dynamical view, if you look at the structure of the world over time, it has incredibly redundant structure. In fact, one might also almost make the conjecture that the basic ontology that we always come up with, the basic ontology of objects, actions, attributes, is an inevitable consequence of the structure of the world. It sort of couldn't help be that way. And so it gives me hope that robots might be expected to develop pretty much the same ontologies as we do. And actually I have some experimental results about that for people who want to talk about it afterwards.

Another reason to hope that we might have both shared meaning and semantic autonomy is that if you embed an agent into a flux where it's got perception, and reasoning, and action—and this is a theme that we've heard many times over the last couple of days—then you are inevitably learning rules that condition action on the perceived state. That's what we do over and over again. It doesn't matter whether you're doing reinforcement learning, or cased-based learning, or some other kind of thing. As long as you have this loop, perception, reasoning, action, it will be necessary to learn these conditional relationships.

And so I think the assertion there is that we're much more likely to have shared meaning and semantic autonomy if we have shared experiences than if we don't. So I really do think that the people at this workshop who have been saying, "Get your agents out there in the world doing things" have a reason to say that. Shared experience leads to shared meaning.

Another theme that's come up in the workshop that is certainly relevant to this tension between semantic autonomy and shared meaning has to do with the deepening, or elaboration, or enrichment of knowledge. I mentioned an idea that Michael and I developed yesterday about how that elaboration might happen. But also let me refer to Ed's notion of progressive tests. You know, you take the test after a year, you take the test a year later, and you're expected to have a richer understanding of your domain when that happens.

We've heard at this workshop about the social environment and the importance of collaborative tasks. And let me just point out that we all learn most of what we know from people, or books written by people. So to look at an agent and say, you know, "Go figure it out yourself without any kind of social interaction" is an entirely unreasonable idea. So the people who are talking about social learning here are, I think, on exactly the right track.

Finally, there was some discussion prompted by Ed. And this really surprised me because a few years ago I wrote a textbook on methodology and Ed said, "What are you doing wasting your time on methodology? Methodology? Nobody's interested in that. Do science." So Ed of course, bless his heart, brought up these methodological issues of focus and integration, challenge problems, and Turing tests. Let me say, as I did yesterday, that I agree with both Ron and Ed that focus and integration are a good thing. But I must point out that the Turing tests had an attribute that the Feigenbaum test does not, and it's an important one. The Turing test was a proxy. Turing himself said, "It will impossible for a machine to pass the Turing test and not be able to pass, indefinitely, many other tests."

The Turing test has nothing to do with conversation, and it has very little to do with sort of general knowledge. It had to do with this proxy role. If you can pass this, you can do anything. These tests of partial intelligence don't have that attribute. I would like to endorse, and acknowledge,

and support whoever it was that said earlier—maybe it was Danny—work on complete AI. Work on the whole thing.

Now, we're not going to be able to do the whole thing in one fell swoop. But that's no reason not to work on it, right? That's a methodological point. The second point about the Turing test is the failure of any test of the form "I know it when I see it," and such tests are enormously useful, is not diagnostic. If you don't see it you don't know anything. So the Turing test has a really big failing in that it's uninformative in failure.

What we really need as tests are diagnostic tests where failure does inform progress. We need that because we need sign posts. We need to see how to move forward. The Hori test doesn't do it either, but I just had to point out that it's—at least for me—the first interesting variation I've seen on Turing's puzzle. I mean, the change in perspective in the Hori test from the individual intelligence of the system to its effects on an entire society is a pretty interesting proposal.

Alright. So I'm almost done. The grand challenge that I would like to leave you with is robots that learn human language. The reason that that's a great challenge is that it brings us face to face with the problems of semantic autonomy and shared meaning. It will require us to make significant progress in perceptual learning, sometimes called the sensor-to-symbol problem. It will require us to ask whether the innate representational primitives out of which representations are—more sophisticated representations are built, and what are the syntax and the semantics of those constructions? And if, like me, you believe that language is something that is layered on top of a conceptual system it will bring us face to face with the question, "what is the conceptual system on top of which language is laid? What is the nature of our most primitive knowledge?"

You'll learn to parse, to ground semantics. And there's good work, a lot of good work, on grounding semantics of verbs and prepositions in interactions with the environment. A lot of people are working on robot language learning now. A lot meaning about five—there are a lot, actually, people working on robot language learning, but those who have made good progress on grounding semantics in action.

It will require us to work on social learning. You know the meaning of a word, I don't. How can I learn the meanings of the words by watching what you're doing and listening to what you're saying? It will require us to think about active learning. I think I know what a word means. I think I know how to construct a phrase. But let me try it out. Let me use the word in some context and see whether I get the reaction that I expect.

Necessarily, we will have to think a lot about compositional semantics, the language of thought hypothesis and so on. And then finally, because although language is layered on a conceptual system it also has the ability to modify a conceptual system we get to see how language closes the loop. And that's a very exciting prospect also.

I said yesterday that I think it's important that foundations be philosophical, that we ask, "What are mental states? How can they have meaning? How does meaning influence behavior? How do lexical semantics become associated with those states?" These are all questions that can be addressed effectively within a grand challenge of this talk. And that's all. Thank you.

Session 7 Panel 4   Discussion

Manuela Veloso:

I am actually so in love with robots that I would say that my challenge is to create an Esperanto language that they also understand. So why should they understand my language? Why can't we devise a language that is a new language? You know, Esperanto didn't have much success. We all speak English, right? So Esperanto died in some sense. But it was a good goal because it was like saying like this, "Nobody has the ground truth, the ultimate truth." In the same sense why can't we devise a new language to talk to the robot? I speak five languages, out of which four fluently, really fluently. So in my head I'm always thinking about different things. But, for example, when I came to the United States "touchdown," "inning," were words that had no meaning to me. But then experience gave it meaning. You are learning the human language, it's something in which there's a lot of experience involved, right?

The robot experience is sensory driven. And so the challenge is how to get that meaning. Twenty years in the United States, and I have no idea what's the meaning no matter how much I watch baseball, right? My challenge in some sense is like this, when are we going to accept limitations? That's it.

Paul Cohen:

My robot learned two verbs as synonyms. And those verbs were "raise" and "close." And you, who have a lot of experience with pioneer robots, can tell me why those were synonyms for the pioneer robot.

The reason that my robot learns "raise" and "close" as synonyms is that when it closes its gripper, it raises it. Sometimes a graduate student will say, "Oh look, it raised it's gripper." And sometimes a student will say, "Oh look, it closed it's gripper." And it means the same thing because it relates those two words to the same pattern of behavior. And it takes those two words to be referring to that pattern of behavior.

Now, that's wrong. It's wrong in the sense that "raise" and "close" are not conventionally taken to have the same meaning. It's right in the sense that it has learned the meanings of words, and that relating words to its experience that way, it's got it absolutely right. So when you say, "Look. It depends on your experience," you know. "There's no right answer. It'll sometimes make mistakes." Absolutely, exactly as it should.

Daniel Bobrow:

On the other hand, one of our goals is to not have these worlds be so separate. The point of it is, we human beings are pretty adaptable. We can take on limitations. And we don't do this for engineering reasons. But I think it's at the goal of artificial intelligence is to have creatures which can live in the world with us, not that we have to move into their world on occasion.

Manuela Veloso:

But then you have to realize that we are generous in making ramps for handicapped people. We are generous to put lights so we can read at night. So we have to do something for robots.

Daniel Bobrow:
We will do something for robots. And we can do these things. And I think there are steps along the way in which we can do those things. But accepting limitations is probably not the right thing. Using the limitations as a resource when you are trying to achieve certain goals, that's a different kind of thing.

Session 7   Discussion

Thomas Dietterich:
At the risk of causing another controversy, I would like to advocate small, incremental steps in building artificial intelligence where we are involved directly in real engineering problems. So I'm a little nervous about Paul's learning semantics by robot, or certainly it's predecessor the L-module project of Learn Base, which was 12 or 13 years ago because I feel like you are the one in charge of defining the task rather than there being some customer who's in charge of defining the task. So I think the progress has come in expert systems, in knowledge-based systems, in machine learning by solving other people's problems.

And certainly Robocup—I mean, Manuela didn't define Robocup tasks. That was, I guess in this case Sony is doing. And that gives us the diagnostic feedback that you were talking about, Paul. And I guess there's a risk that we will be too modest in our long-term goals. But there's also kind of the comforting guarantee that we will already be helping society each step along the way. And particularly we will continue having the interest of our funding agencies, which I think is also an important point. So I know we've been thinking big and grand here, but I want to also think incremental and small at the same time if that's possible.

Eric Horvitz:
The concepts that Paul brought up, and that have been brought up by others during this discussion, center on linking language and internal representations to primary sensory and motor information. Others working in this realm include Deb Roy at MIT. There are a number of interesting challenges in developing machinery for constructing, in an autonomous manner, some useful understanding of the world from streams of perceptions, perhaps starting with very primitive axioms about geometry, and so on. Rather than viewing this as a means for directly developing deep intelligence, the goal of this work can be taken as learning about how some intermediate level of competency might evolve or be evolved, and thus serve as a small incremental step in developing autonomous intelligences. So, I don't think we need to take some of the concepts that Paul presented as focused on taking those basic, simple accelerometer readings up to a large-scale macroagent that's doing high-level reasoning. I think that these kinds of pushes at the sensor or motor end of things have not been pursued deeply enough, even if for only building incrementally toward intelligence.

Stuart Russell:
It wasn't clear to me at all what was essential about the robotics domain. I certainly think that language should be situated. But I could perfectly well imagine a much more sophisticated situation without bumping and dropping, and in some sense separate out the robotic sensory issues from the language issues. And I think that we have to be clear about what, precisely, are the research goals and how you go about them, solving those goals.

But the next point that I want to make is that, partly from the experience of writing definition, I've been getting the feeling that AI is in danger of succeeding. Not immediately. Not in the next five years, but I think there is now a much more serious danger of success than there was ten years ago. But I also think that there's a danger of AI failing. And I think that one way that could happen is by the potentially collective failure of the field to develop the kind of, shall we say

community infrastructure that's needed to maintain a successful research enterprise on the scale that we expect to need if we are to develop more intelligent systems. We don't have the kind of infrastructure that other engineering societies have in terms of training programs, the development of standards, large scale masters programs or training programs backed up by very large industries that rely on all these structural developments. So I think that that's a job for the professional societies. AAAI does not have an education program. I think that it's a disgrace, and I blame it partly on myself. As I said, I would try and get it going.

Michael Witbrock:
I heard a lot of interesting work, and there seems to be a general desire to work together to produce something more in the way of AI. And I think it would be a great pity if we left here without some form of concrete collaboration coming out. So I, in particular, I heard some very interesting things about robot sensing, and I think that there's a possibility for collaboration. We heard some very interesting examples of natural language processing where in particular I think Cycorp and I could do some worthwhile collaboration. So from my point, I'd like very strongly to encourage people to contact me if there's any way in which you can collaborate. And since there are about 40 of us here, I hope there will be approximately 1600 or 800 such interactions going on following up from this meeting.

Hiroki Arimura:
I want to add one small thing to the language problem. The missing link is, I think, the syntax. There are many discussions about the emergence of syntax, Chomskian vs. non-Chomskian view. And the semantics, there was great discussion about it. But still many people, I think, believe that syntax is only of human species. And so probably intelligence, and one of the reasons that we have high intelligence in humans, is that humans have syntaxing. So the syntax could be a basis of not only logics or mathematics, but also arithmetic and other common beliefs. So I think one, maybe small thing—or maybe large thing—is to study syntax.

Paul Cohen:
The idea of learning language by being embedded in the environment is a good idea that is showing good results. And it isn't necessary that that environment be a physical environment. But there are lots of reasons I want it to be. One is that's how my daughter developed, and I entrusted in developmental psychology. Another is that the semantics of verbs and prepositions are being explained well by linguists, not computer scientists but linguists, by relating to physical parameters. But the other linguists are largely concerned with syntax. So there are two schools of linguists. One school of linguists is concerned with syntax. And the other one is concerned with explaining some syntactic conundrums in terms of semantics that ground out the physical structure of the world.

So I think it's a good thing to try to learn a language in the context of physical activity. I think it's turning out to be very productive. To Tom I would say that it's necessary so that my students can get jobs that we make many small incremental steps. And so we develop algorithms. And the stuff that I showed you was an extremely small step. But the nice thing is that after that step, if you look at this in the most generous way, then a robot has actually learned the meaning of the word for itself.

Thomas Dietterich:
People in pattern recognition have been learning this for years. I've got a computer system that knows what *ascer scircinathine* means, right? Because that's a particular species of tree.

Manuela Veloso:
But I think there is something here that Eric talked about, and I think Paul has mentioned it already. Maybe people have been processing signals. I mean, signal processing is a field that's old. However, trying to go from the signal all the way to actual meaning and, for example, when I do the HMM's actual behavior like sequences. So that is kind of the AI job. The goal of signal processing is too disconnected from the actual embedment in some behavior, or in some meaning, or in some activity of the user. So that's actually something that we kind of have to fill that gap. Of course there are people that already did signal processing, but never with the AI goal in mind. And that's where we are.

Stuart Russell:
But they're recognizing speech. They're recognizing several thousand different kinds of events, as opposed to just 10 or 20.

Manuela Veloso:
But they only relate to one kind of sensory input, which is the recorded acoustic signal. They are not learning any meaning.

Daniel Bobrow:
It's not the difference between large steps, one large step and one small step. It's whether you have a staircase. And part of the problem is not whether you can pretend a step to be a whole bunch of stairs but whether there's a sequence of these things. I mean, what's impressive about what Manuela has done is not that she's gotten the robots to where they are. But there's been a sequence of these things, an on-going set of work that keeps getting better. The fear is, saying it's robots and that's the only place to do it. And so that's the only staircase up to where we're going.

So I think we need to do small experiments to actually learn. We have to know what we're trying to learn. And then we have to understand how we're going to build on that. Or, what we see time and again—I've seen this since the very beginning—is a little step squinted at looks like you are going to solve the rest of it. I remember when we had the first perceptron distinguish between o and x, we thought we had the pattern recognition problem solved. So I think we have to understand what we've learned, and we have to build on that.

Epilogue

The Future of AI workshop has become a reality after nearly one year of preparation. It was a wonderful workshop. The 30 foremost AI researchers gathered from Japan (20) and United States (10) to Amagi and spent two days to discuss various issues of AI research based on their own experience. This was the first attempt in such a meeting. Everybody enjoyed the presentation and the discussion that followed thereafter. The topics covered a variety of important research areas. We visualized from the very beginning that there will be no complete consensus, nor concrete conclusion; rather we thought that it is important to be aware and to admit that people see things differently and to argue why they think this way, not that way.

This volume offers a summary of what we discussed during these two days. We believe that readers will find the messages delivered in this volume valuable to their views of AI research. All the participants felt that it is worthwhile to continue to have this series of workshop in the future. We plan to organize the second workshop in two years from now, probably in Japan. Nothing concrete has been decided yet, though.

The contents of this volume have been prepared with much care, but the discussion transcriptions may be less than perfect in its integrity due to the low quality of the recording caused by unforeseen hardware equipment trouble. We hope the readers will find that minor technicalities do not take away much from the true value of the contents.

We wish to express our sincere thanks to all the participants for accepting our invitation and contributing to the success of this workshop. We also extend our thanks to all the sponsors for their financial support without which organizing this workshop would have been impossible.

July, 2003

Future of AI Workshop
Steering Committee Members
  Edward A. Feigenbaum
  Setsuo Ohsuga
  Hiroshi Motoda
  Koji Sasaki