
Welcome
2017 Faculty Summit
Attendees
Faculty Summit 2017
microsoftfacultysummit.com
Microsoft Research
Microsoft.com/research
Facebook
@microsoftresearch
Twitter
@MSFTResearch
#FacSumm
#EdgeofAI
Microsoft Research
Faculty Summit 2017

Microsoft Research
Faculty Summit
2017

Robotics and AI

From Artificial Intelligence to Augmented Intelligence

Katsu Ikeuchi
MSRA – Redmond, Microsoft

The beginning of AI

- The ***Dartmouth conference*** in 1956


McCarty


Minsky


Rochester


Shannon

...

- ***Robotics*** and ***Computer vision*** were synonyms with ***AI*** around that time

- In fact, the artificial intelligence laboratory, MIT, one of a few AI center at that time, completed the following demos:

Copy Demo @The AI Lab, MIT

15 years after the Dartmouth

- **Sensing**: to watch the block world
- **Cognition**: to understand the structure
- **Action**: to create the same structure

Target scene

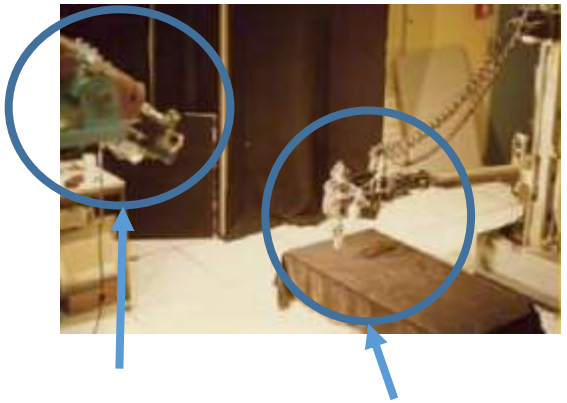


Image dissector for sensing

Manipulator for action

Binford (S)

Winston (C)

Horn (A,S)

Minsky

Bin-picking Demo @The AI Lab, MIT

25 years after the Dartmouth

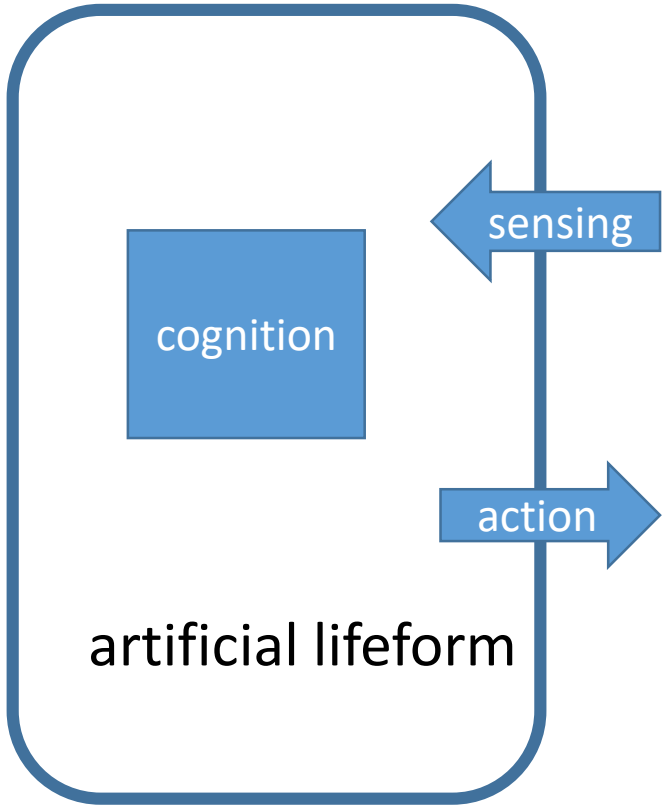
- **Sensing**: to watch a bin through photometric stereo
- **Cognition**: to understand the configurations through Extended Gaussian Image (orientation histogram)
- **Action**: to grasp the top-most object based on the obtained configuration

Horn & Ikeuchi *Scientific American 84*

The goal of AI (Robotics and CV) at that time

- To create an artificial lifeform (robot) with the three components:
  - *Sensing*
  - *Cognition*
  - *Action*

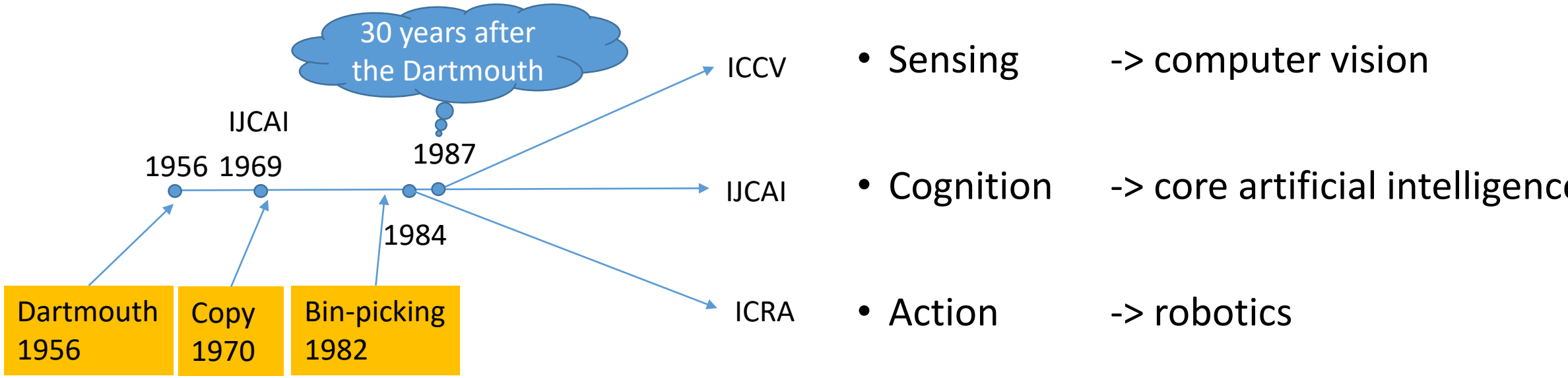
- The demos
  - *Automatic* and *autonomous* systems
  - *Static environment* without human intervention and human interaction



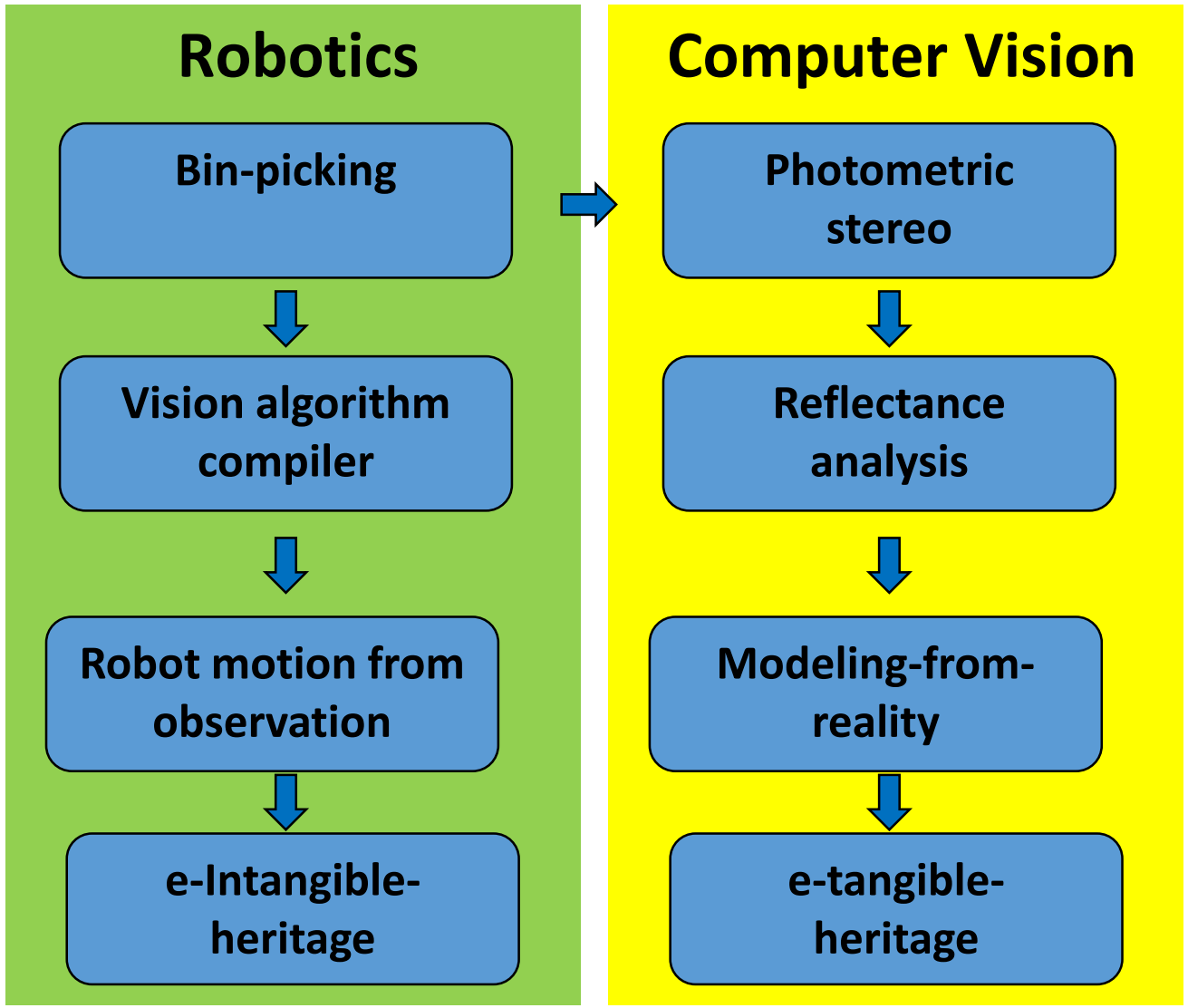
Still, the three components contain challenges

Reductionism!

- System -> component
- Divide-and-conquer



My personal history along the reductionism



Reductionism!

- System -> component
- Divide-and-conquer

The main issue in the bin-picking demo

- The system looks intelligent!!

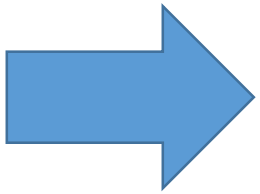
- But, Ikeuchi (*a programmer*) considers all the conditions and programs the robot before hand

- TELE-OPERATION through *the program by Ikeuchi (a programmer)!*

- Can a robot automatically generate such programs to control the robot?

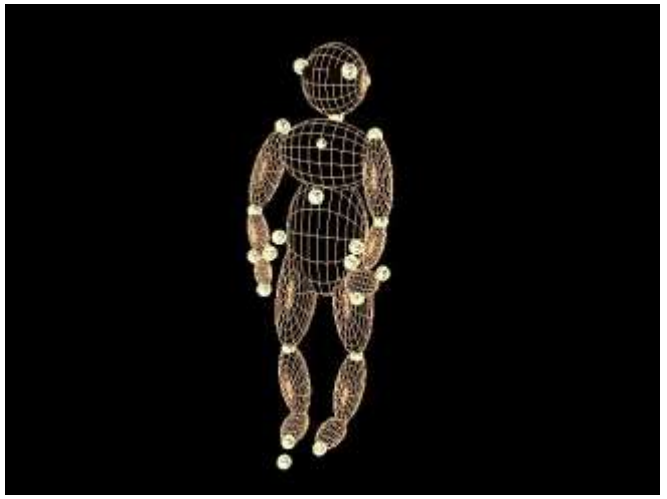
If a robot can learn how to do from observing human performance, → then, the robot is intelligent

Observation



Performance

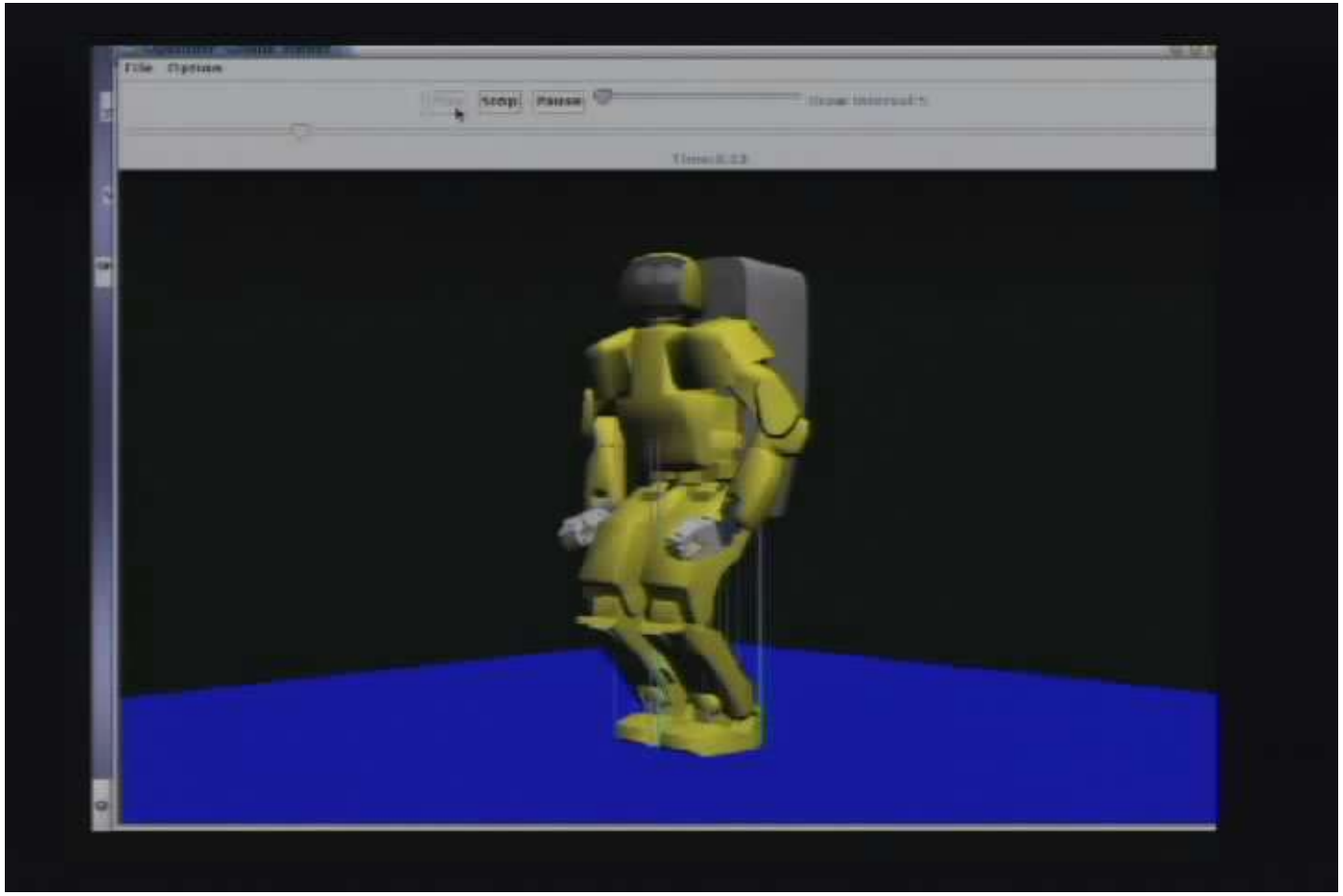
Let’s suppose we have perfect date from observation



Human Performance

Motion capture date
(noise-free, no occlusion)

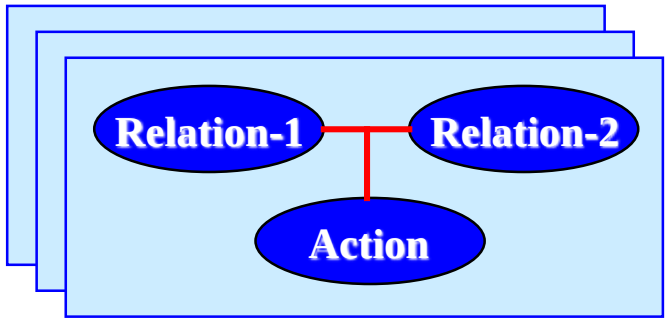
Without understanding (direct mimicking)



AIST dynamic simulator

Need to know the key essences

Observation



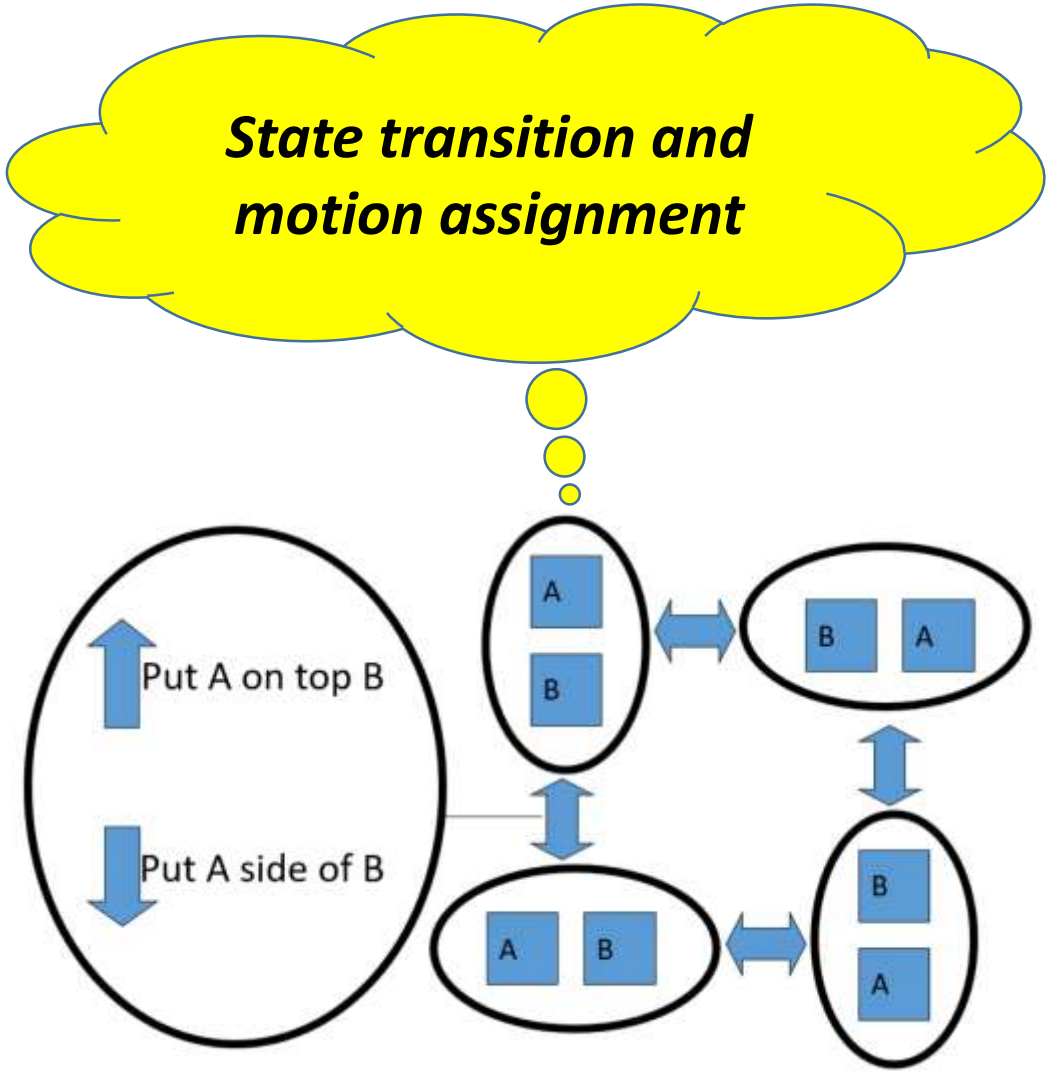
Understanding
the key essences

Direct mimicking

Performance

Learning-from-observation Demo@RI, CMU

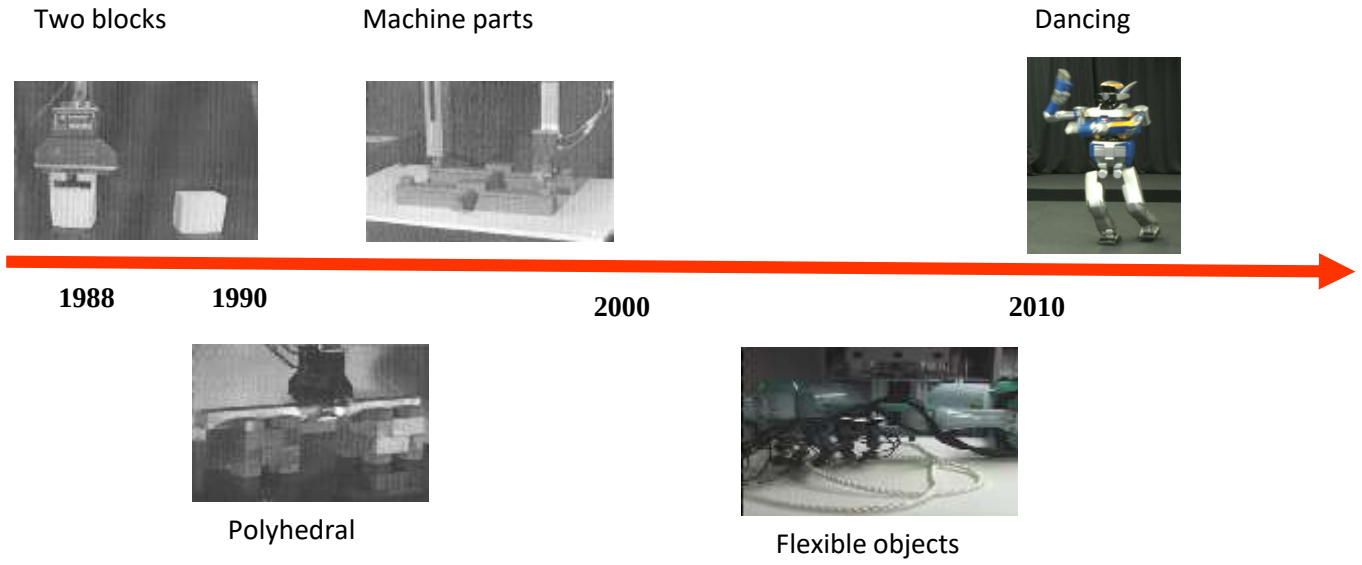
30 years after the Dartmouth



Ikeuchi and Reddy, Robotics review, 1988

Divide and Conquer

- Operation (contact transition) is the main goal
  - Polyhedral objects
  - Machine parts
  - Flexible objects
- Gesture (pose transition) is the main goal
  - Folk dance

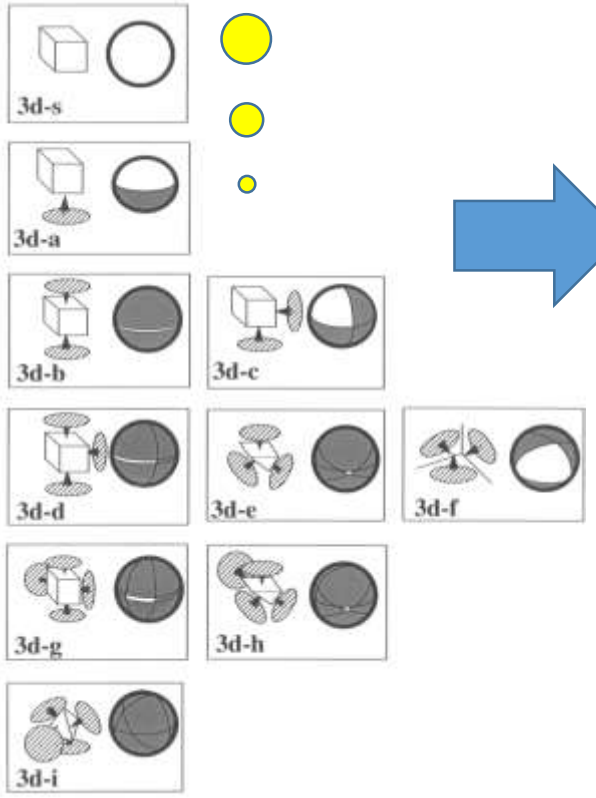


Polyhedral objects @RI, CMU & UTokyo

40 years after the Dartmouth

Kuhn Tucker Theory
face contact among
polyhedral objects

Human demonstration



Suehiro&Ikeuchi *IEEE-TRAS 1994*

Robot performance

Takamatsu et. al. *IJRR 2002*

Daily-life objects (no-contact) @UTokyo

50 years after the Dartmouth

Human demonstration

Trajectory integration

Ogawara et. al. *T-IE 2003*

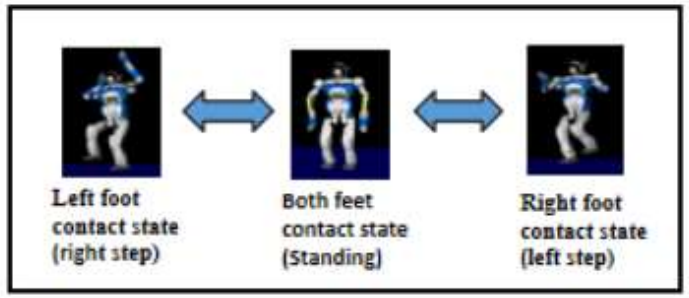
Robot performance

Body motion @ UTokyo

Lower body rep
Foot contact relation

50 years later after the Dartmouth

Human demonstration



Nakaoka et al. *IJRR 2007*

Robot performance

Upper body rep
Joint angle

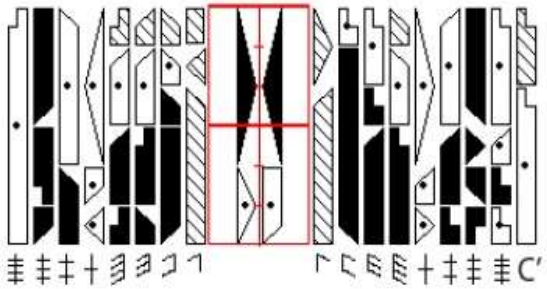
Body motion @U Tokyo

50 years later after the Dartmouth

Human demonstration

Upper body rep

Labanotation
representing human gestures



Okamoto et al *IROS 2010*

Robot performance

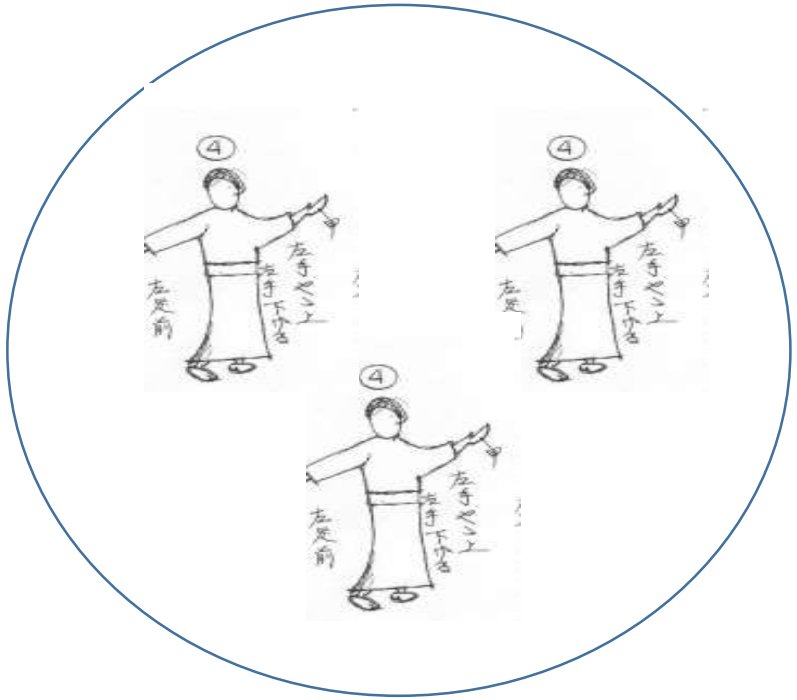
Further detour from robotics

Labanotation and dance analysis

Same dance?

Same dance?

How to define the same dance set

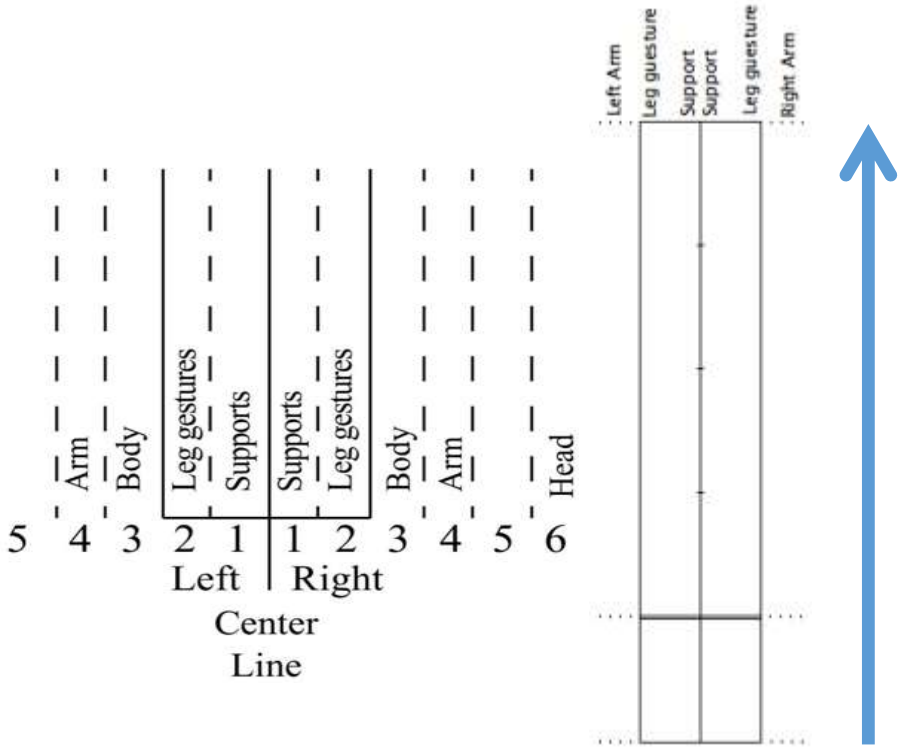


Different dance

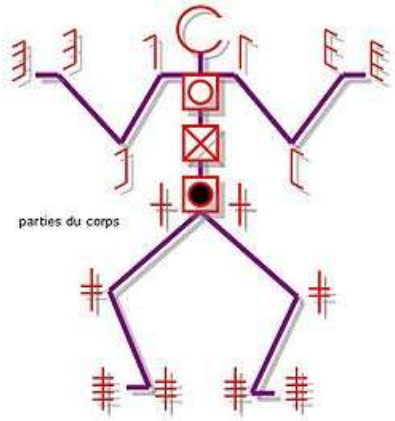
Same dance set

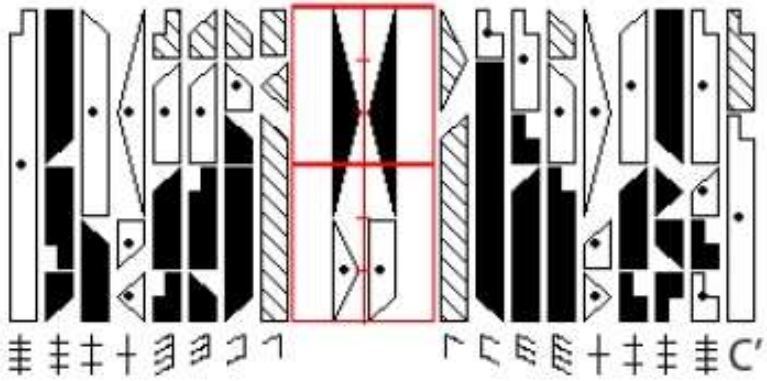
Those dances are same in terms of human perception

Labanotation



R. V. Laban 1927





Labanotation Score

same in human perception

Necessary condition
Sufficient condition

Labanotation

Labanotation

same in Labanotation

Recording Taiwanese indigenous dance based on Labanotation

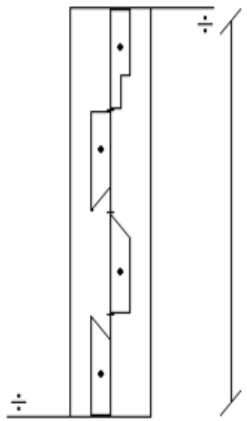
Why Taiwanese Indigenous dance ?

- Considered as the origin of Austronesian
- First to settle on the Taiwanese island
- 14 tribes identified
- Lost culture along modernization due to no-written symbols

Paiwan dance and Labanotation

4-Step Dance

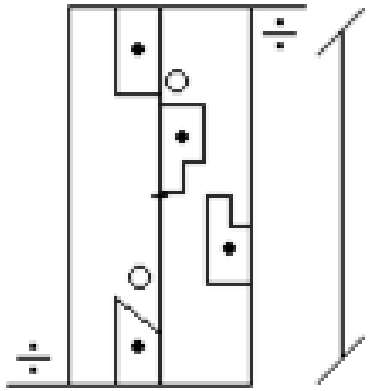
[redacted]



Labanotation of
Paiwan 4 step dance

2-Step Dance

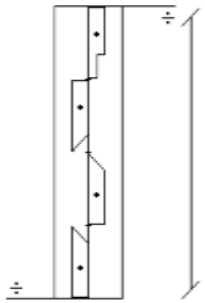
[redacted]



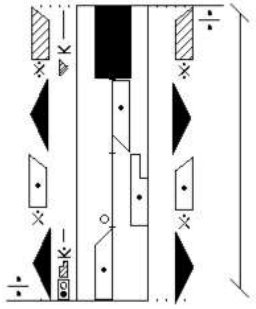
Labanotation of
Paiwan 2 step dance

4-Step Dance

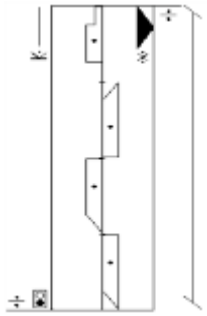
Rukai



Kavalan

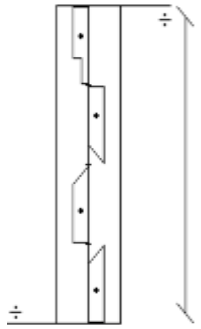


Sakizaya



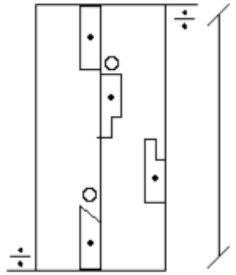
Puyma

Tsou

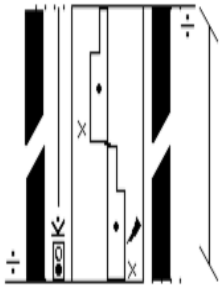


2-Step Dance

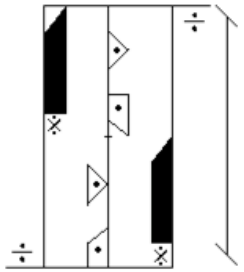
Rukai



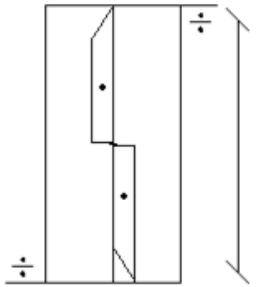
Kavalan



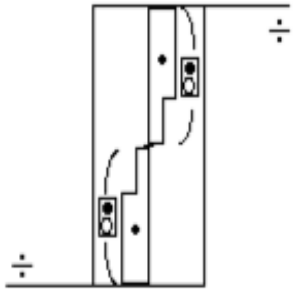
Sakizaya



Puyma



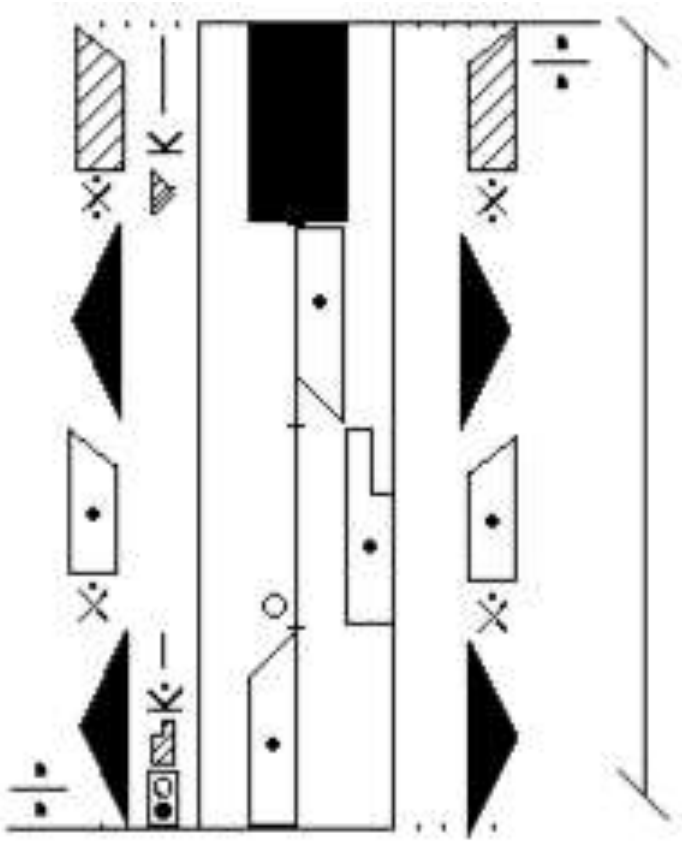
Tsou

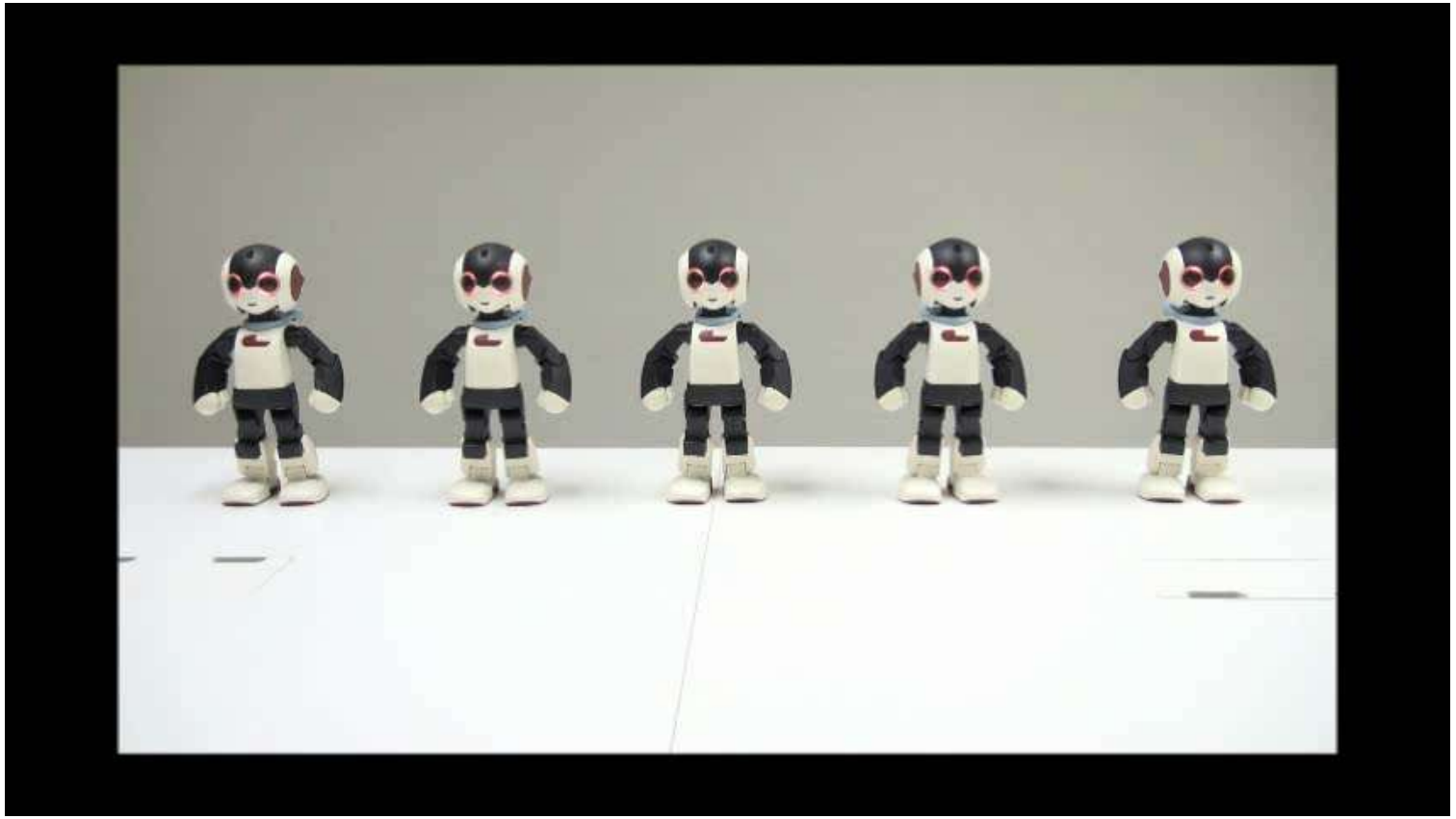


Summary of festival dances

No.	Group	4-Step	Direction	2-Step	Direction
1	**Amis**	Y	Both	Y	Both
2	**Kavalan**	Y	CCW	Y	CCW
3	**Sakizaya**	Y	CCW	Y	CCW
4	**Puyuma**	Y	CCW	Y	CCW
5	**Paiwan**	Y	CW	Y	CW
6	**Rukai**	Y	CW	Y	CW
7	**Bunun**	N	NA	N	NA
8	**Atayal**	N	NA	Y	CCW
9	**Seediq**	N	NA	Y	CCW
10	**Teuko**	N	NA	Y	CCW
11	**Saisiyat**	N	NA	Y	CCW
12	**Tsou**	Y	CCW	Y	No move
13	**Thao**	N	NA	Y	Both
14	**Tao (Yami)**	N	NA	N	NA

Labanotation of Kavalan dance and reconstruction

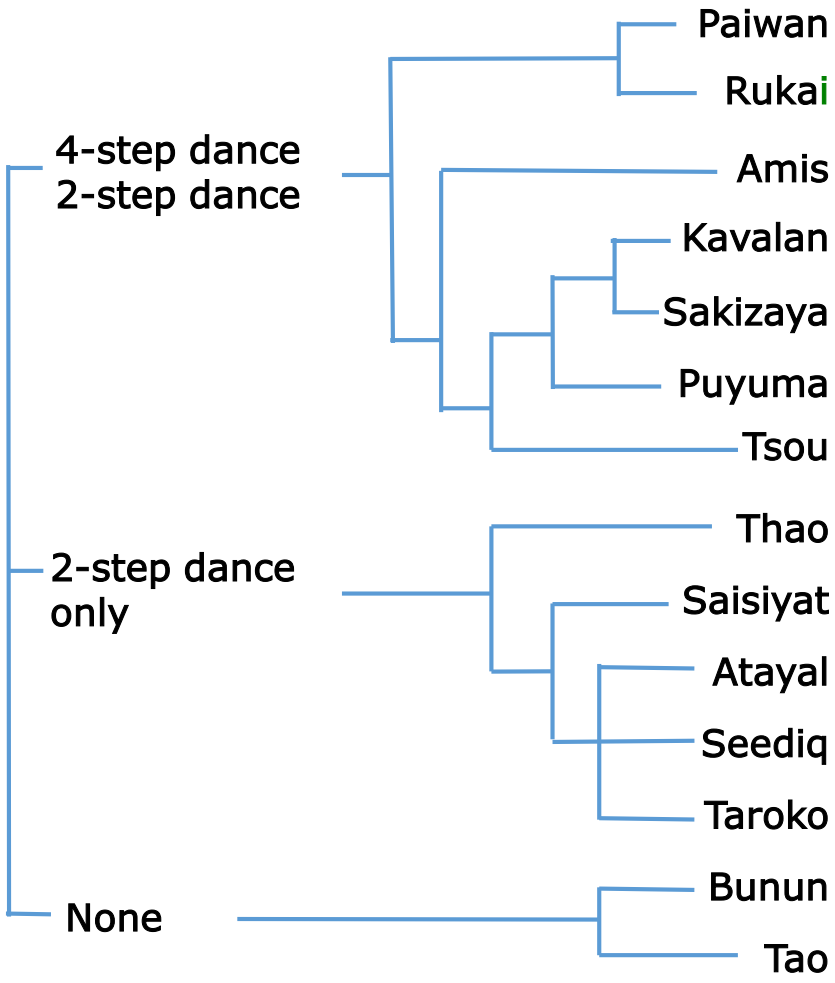




© Prof. Takahashi (a Robot Creator)

Analyzing Taiwanese indigenous dance based on Labanotation

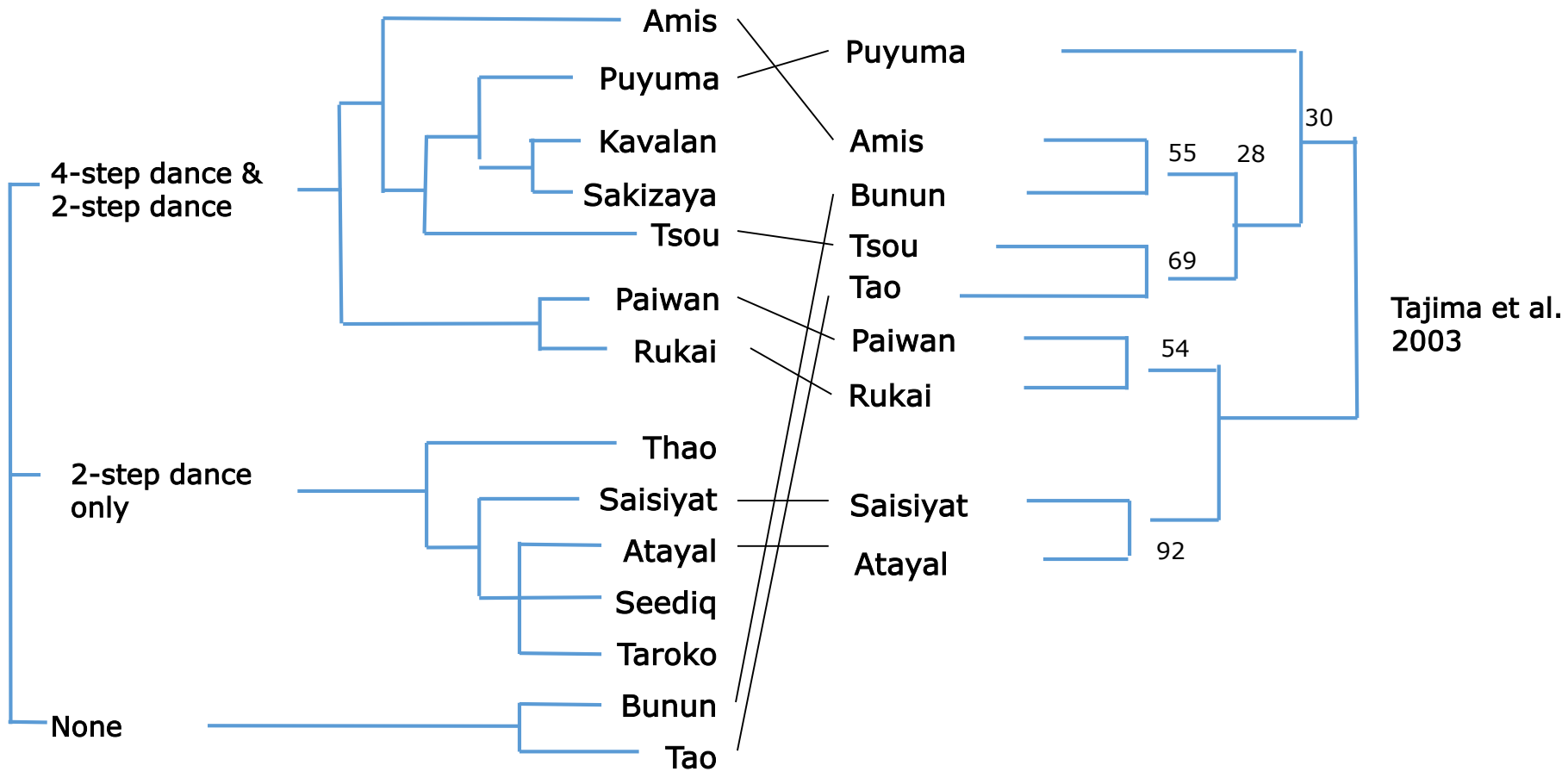
Classification of dances



	4-step	2-step	
Paiwan	CW	CW	(Parallel foot)
Rukai	CW	CW	(Left foot in front)
Amis	BOTH	BOTH	
Kavalan	CCW	CCW	(Left foot in front)
Sakizaya	CCW	CCW	(Left foot in front)
Puyuma	CCW	CCW	(Right foot in front)
Tsou	CCW	NO	(Left foot in front)
Thao		BOTH	
Saisiyat		CCW	(Right foot in front)
Atayal		CCW	(Right foot in front)
Seediq		CCW	(Right foot in front)
Taroko		CCW	(Right foot in front)

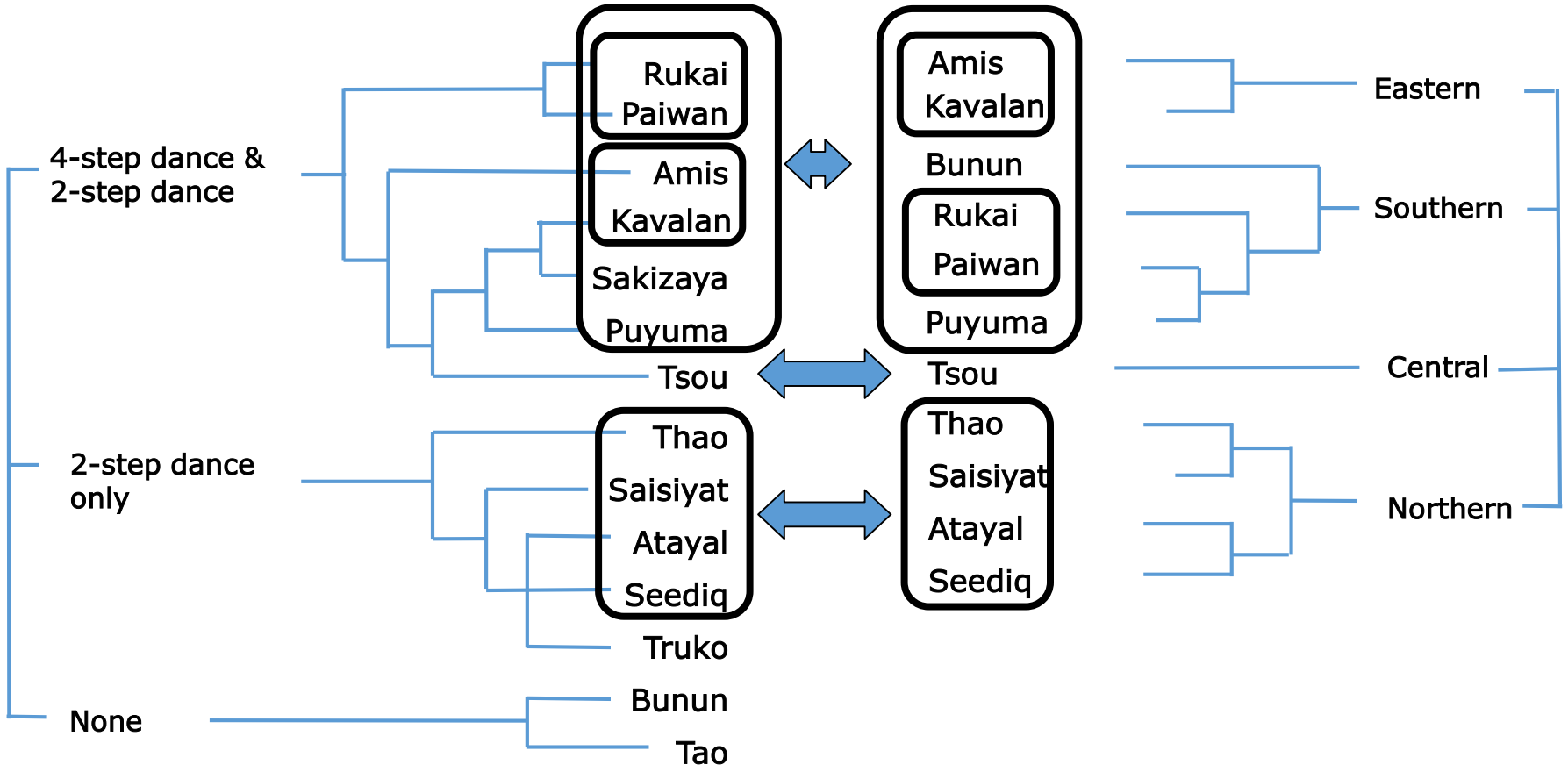
Hu et.al. *CC 2014*

Comparison with DNA-based dendrogram (Tajima's analysis)



Comparison with Dialects

Hu et.al. 2014

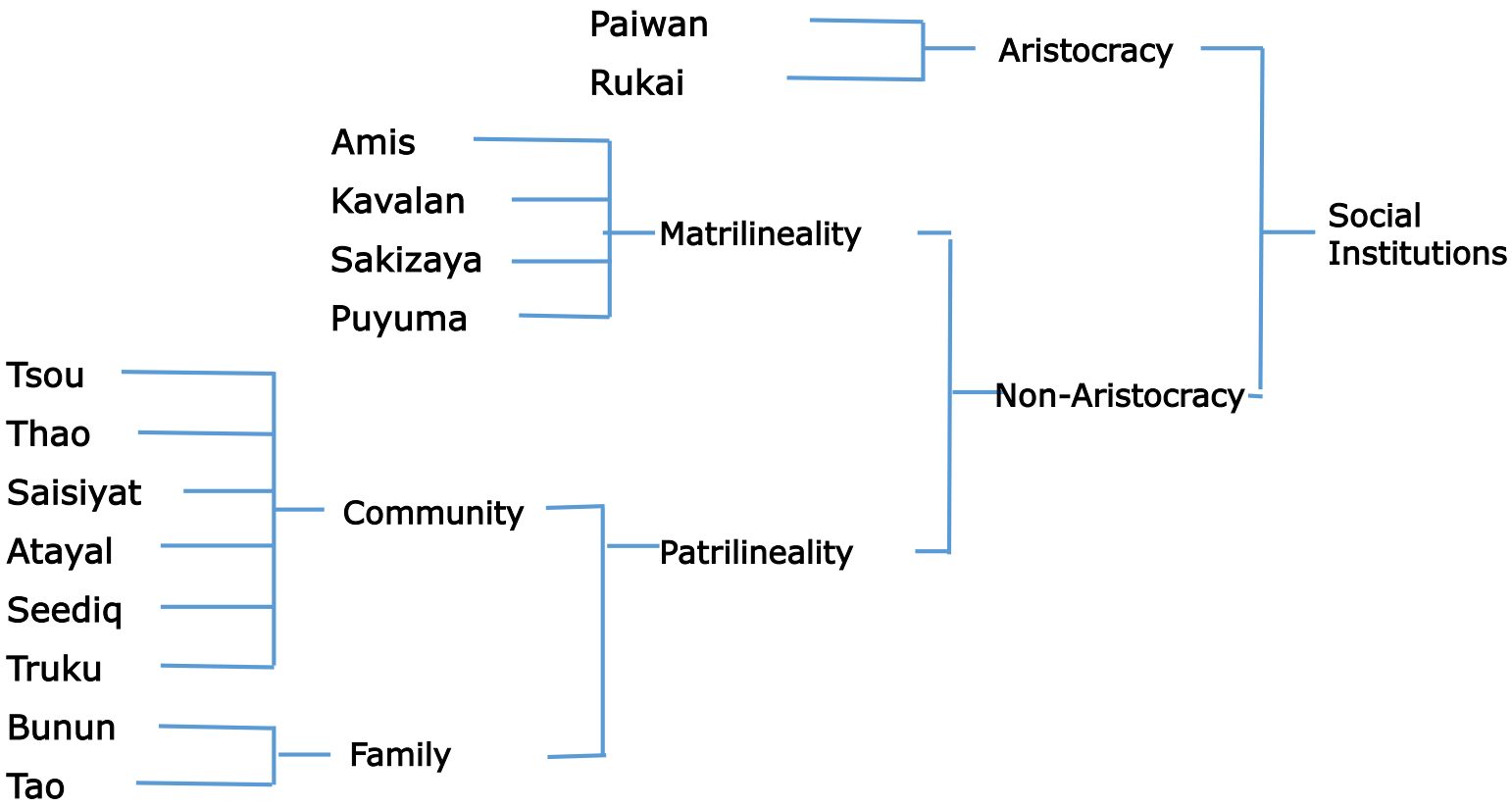


Social Institutions

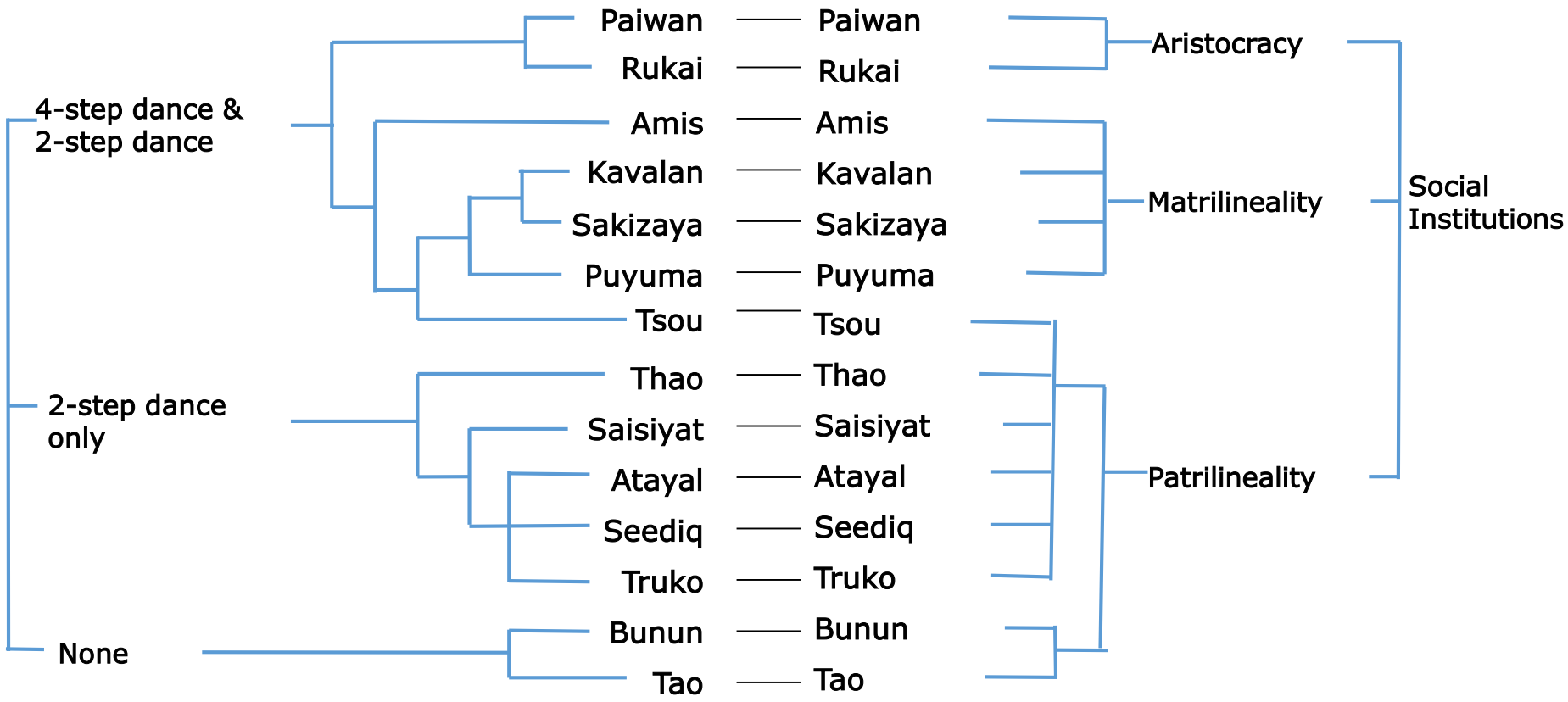
- **Inheritance in Social Position**
  - **Aristocratic:** leader, nobility, warrior, civilian
  - **Non-Aristocratic**
- **Non-Aristocratic: Inheritance of assets**
  - **Matrilineality: mother side**/the eldest daughter
  - **Patrilineality: Father side/**the eldest son

    Men's position are usually higher than females

Social Institutions



Comparison with Social Institutions



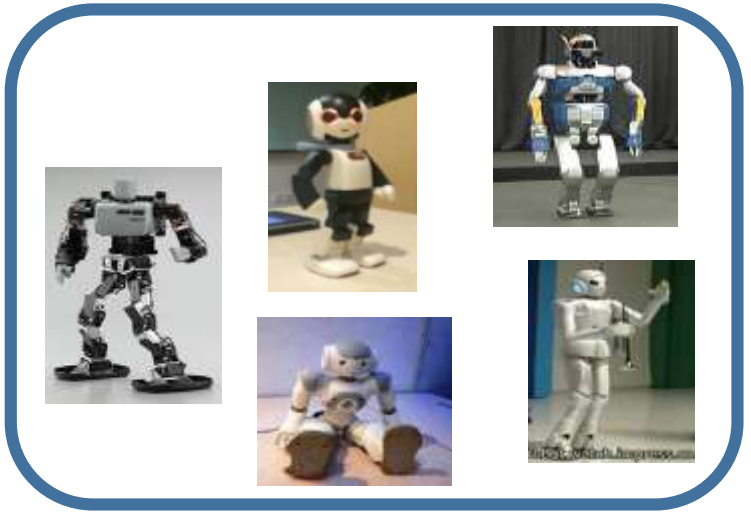
Short message

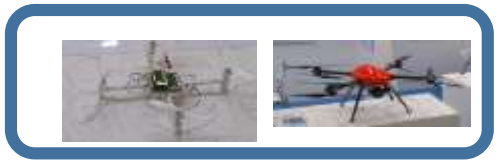
- Folk dance and society structure are closely related
- CS technology can help preserve cultural heritage
- CS technology can help some other scientific discipline → cyber humanities

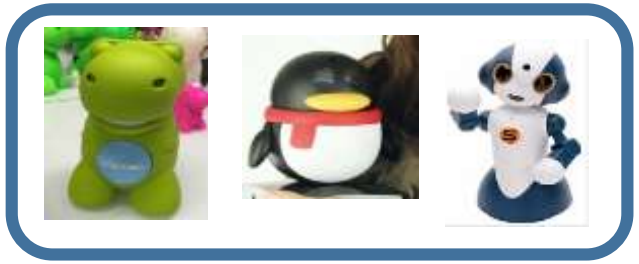
Back to Business

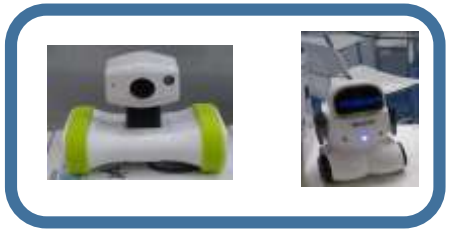
The Cambrian explosion in Robotics

60 years after the Dartmouth

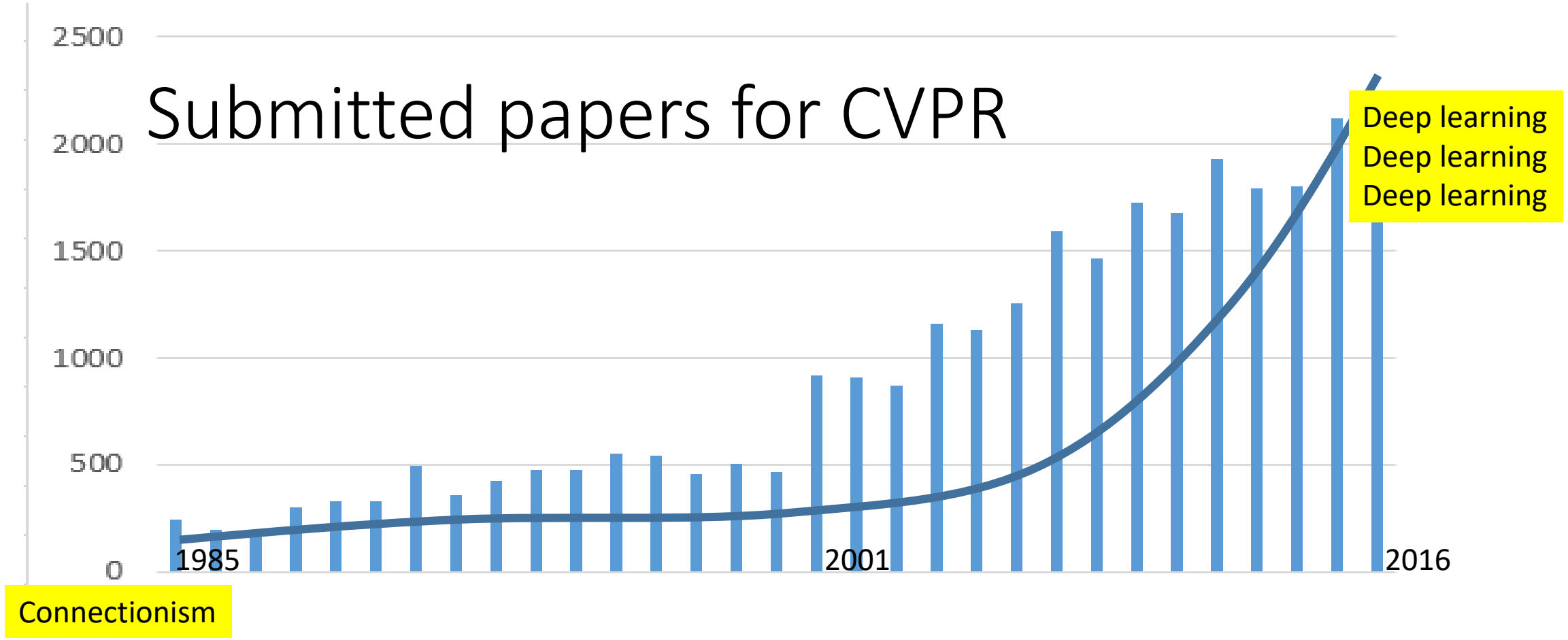






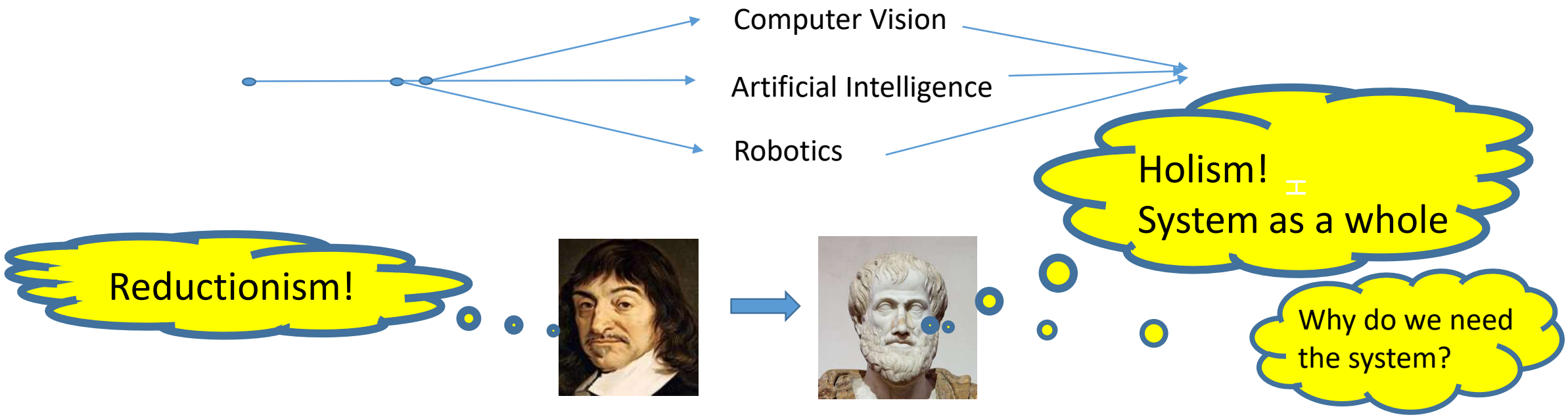


The Cambrian explosion in Computer Vision



Need paradigm shift

- Too much reductionism -> disciplines are too fragmented
- To introduce Holism approach to foresee as a system with a clear goal

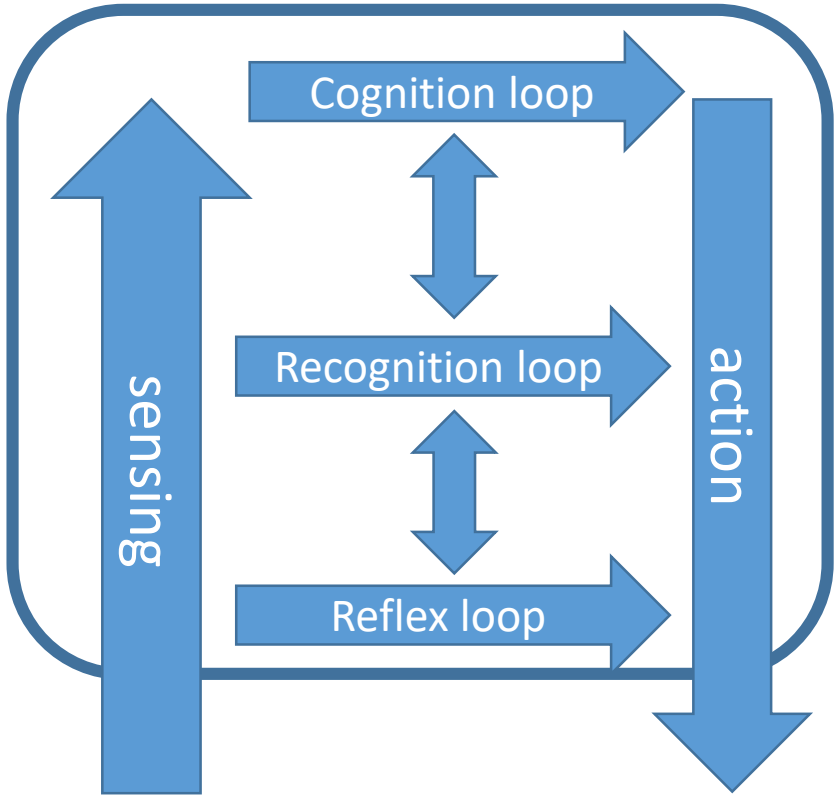


Holism and System architecture

- Cognition loop : Human robot interaction

- Recognition loop: manipulation, navigation

- Reflex loop: obstacle avoidance, balance maintenance



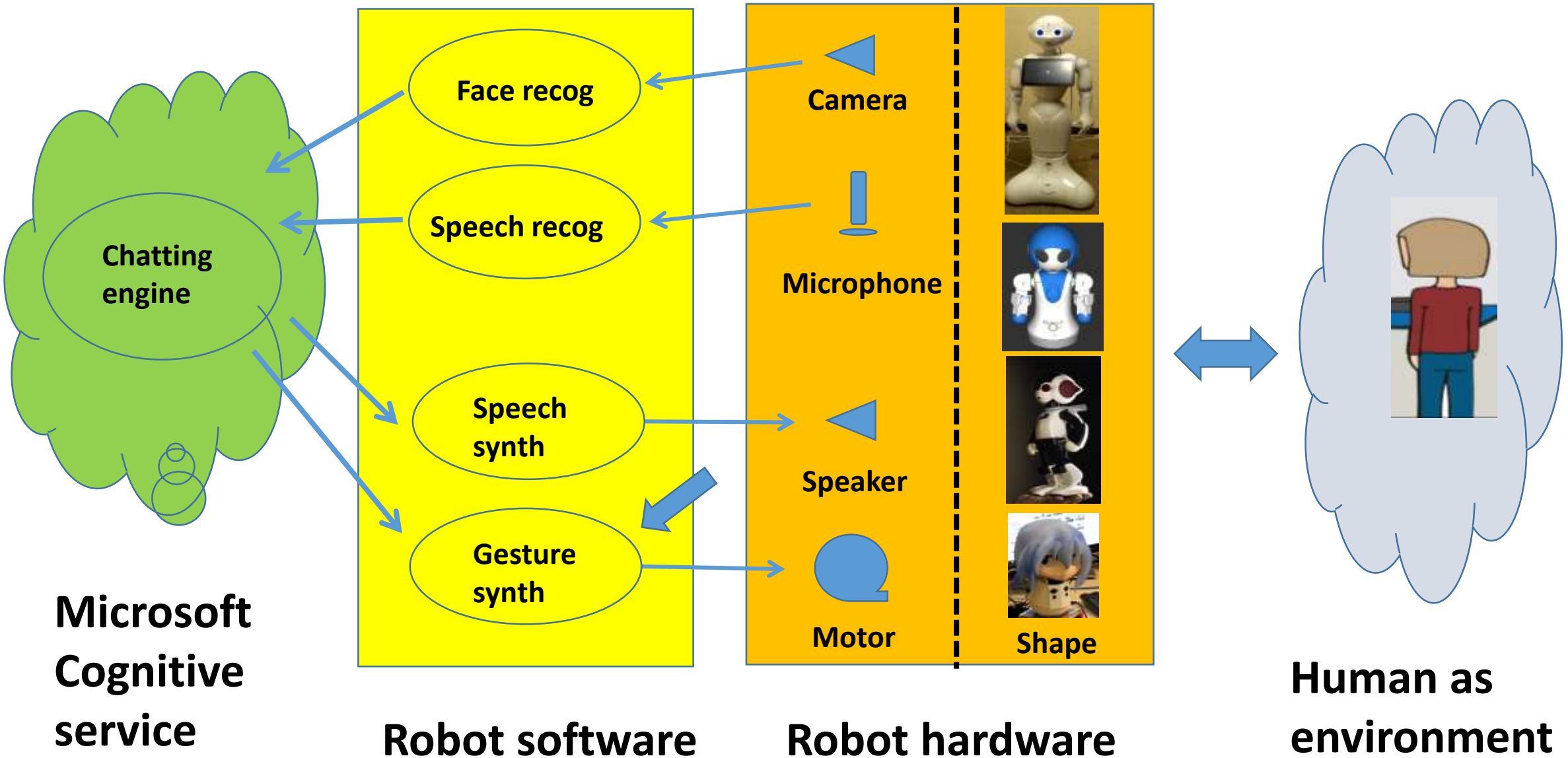
- Brooks' subsumption architecture

Brooks *IEEE JRA 1986*

Closed loop systems in my system

- **reflex**
  - Balance maintenance（dancing robots）
  - Proximity sensing（bin-picking）
- **recognition**
  - Recognizing Human gestures（dancing robot）
  - Configuration of objects to be grasped（bin-picking）
- **cognition**
  - Missing loop in my endeavor

Cognition loop with machine independent software

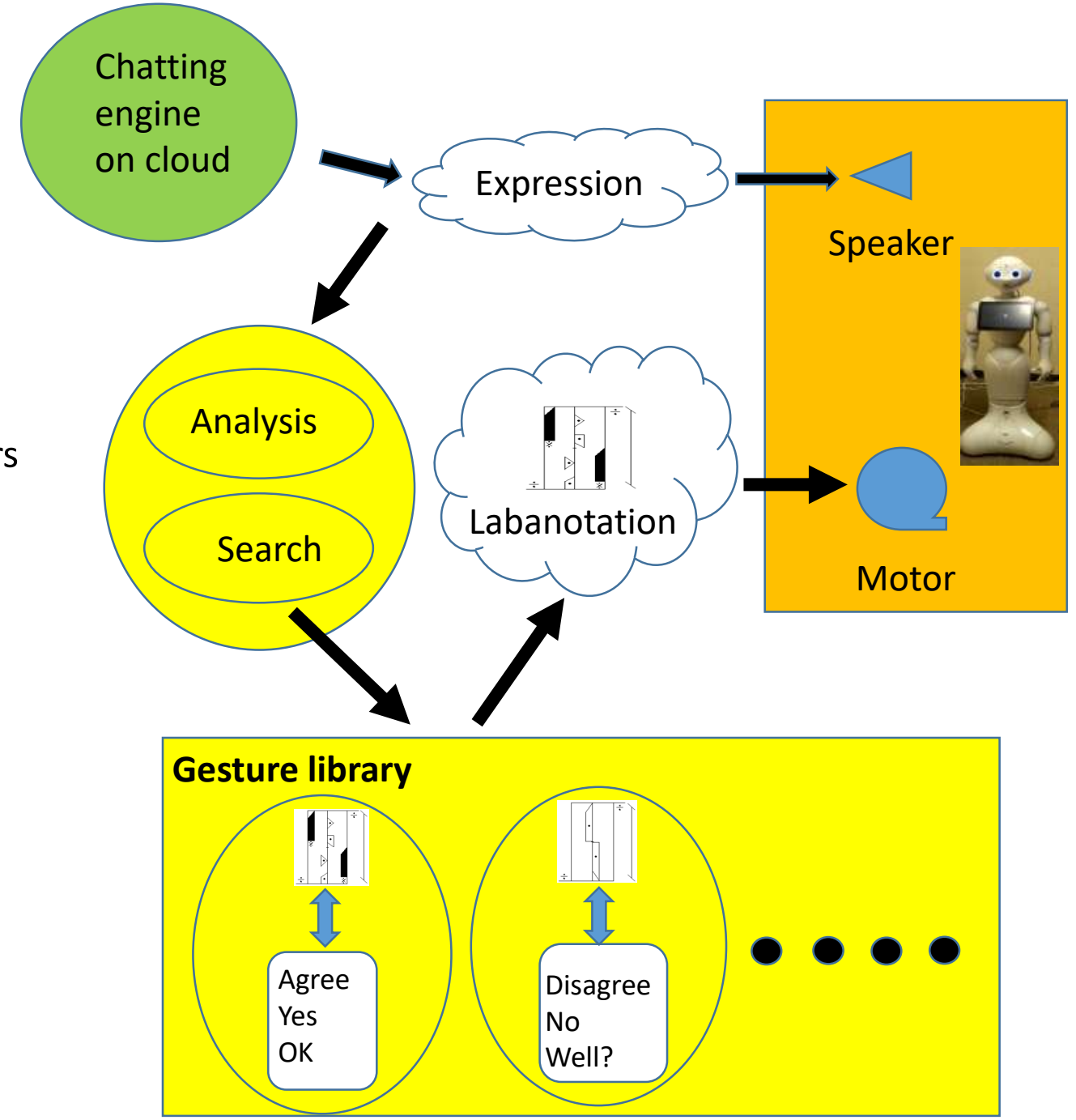


Gesture synthesis

- **Translation of spoken language into body language**

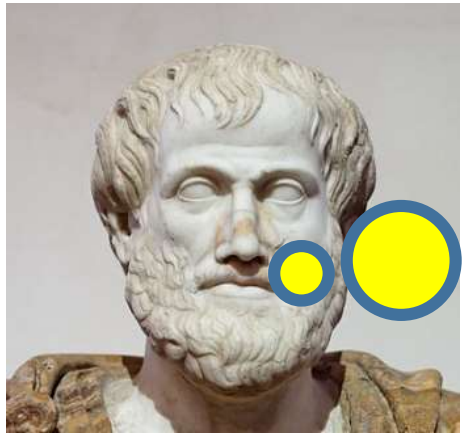
- **Off-line**
  - Gesture library: Reservoir of concept-gesture pairs
  - Gestures represented as Labanotation

- **On-line**
  - Chatting engine generates an expression
  - Analysis of expression
  - Search appropriate gesture
  - Convert Labanotation into motor control signals



Chatting robot with behavior learning

- **Reflex loop**:
  - motion control
- **Recognition loop:**
  - Recognize human face
  - Recognize human gesture
  - Recognize objects
  - Recognize human speech
- **Cognition loop:**
  - Question and answer
  - Search in the web
  - Accumulate gestures
  - Generate behavior library

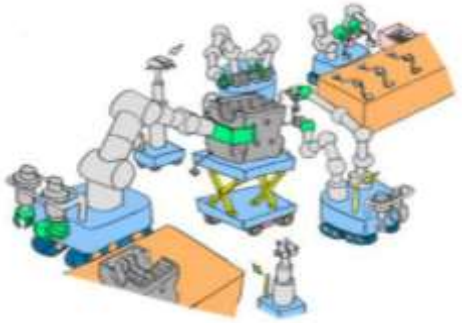


- The current chatting robot is just a composition of components
- The robot system is not optimized for the environment

-> Define the environment!

Environments = Application areas

- **Industrial robotics** (indoor/clean/pre-arrangable)
  - Assembly/Disassembly of industrial products
  - Logistics – sorting objects and picking-placing products



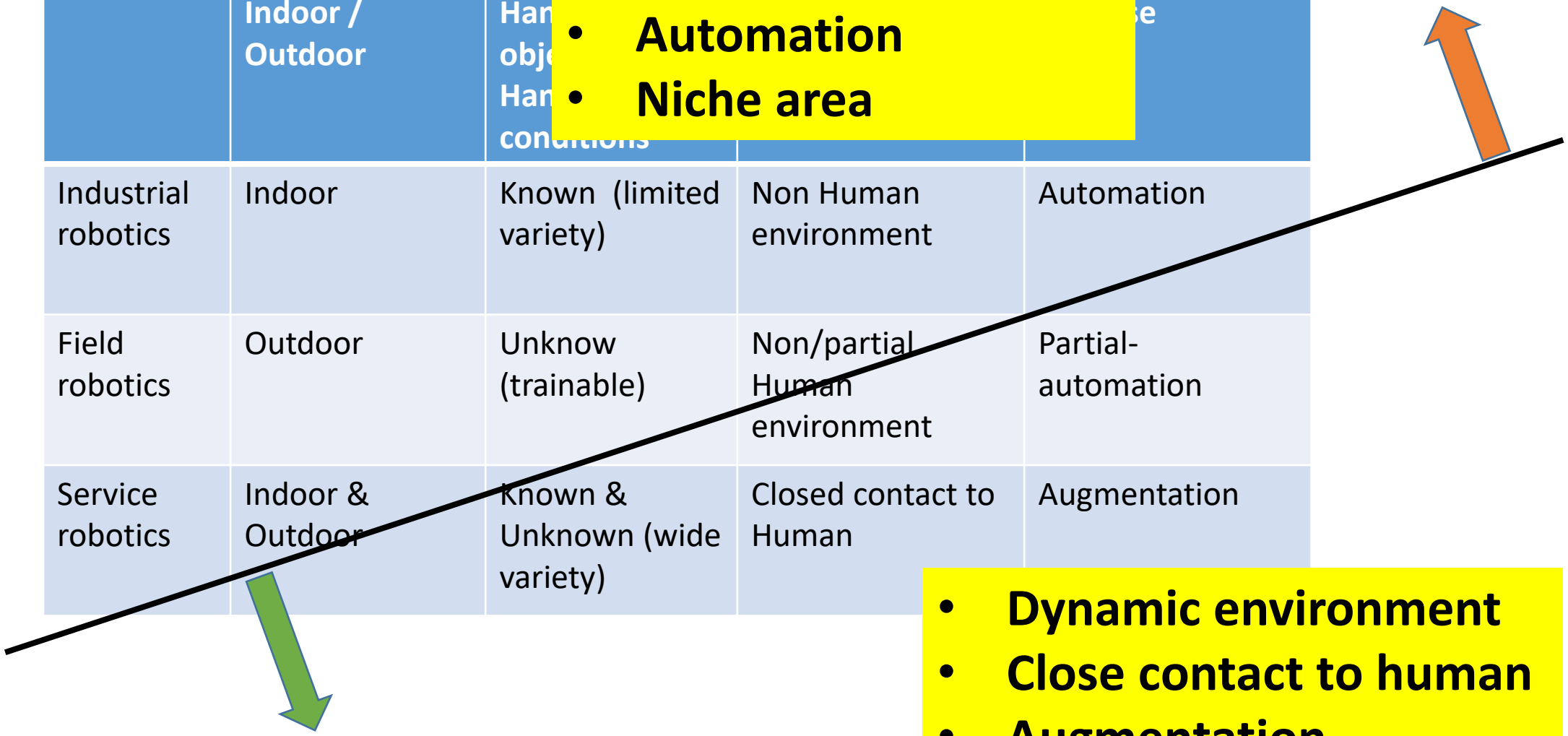
- **Field robotics** (outdoor/random/un-expectable)
  - Plant disaster response
  - Tunnel disaster response
  - Defense response

- **Service robotics** (indoor-outdoor/random/human)
  - Home service
  - Enterprise service

- **Static environment**
- **No human**
- **Automation**
- **Niche area**

Old challenges

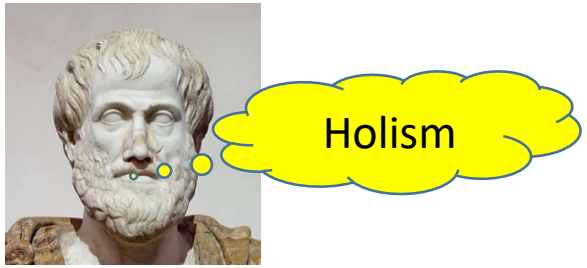
	Indoor / Outdoor	Han obje Han conditions		se
Industrial robotics	Indoor	Known (limited variety)	Non Human environment	Automation
Field robotics	Outdoor	Unknow (trainable)	Non/partial Human environment	Partial-automation
Service robotics	Indoor & Outdoor	Known & Unknown (wide variety)	Closed contact to Human	Augmentation



New challenges

- **Dynamic environment**
- **Close contact to human**
- **Augmentation**
- **Broader area**

Home environment



- Goal
  - Elderly care / Household support

- Required recognition loop
  - Face and gesture recognition tuned to the user
  - Conversation tuned to the user
  - Situation recognition tuned to the user

- Required cognitive loop
  - To memorize daily events
  - To physically help some actions
  - If necessary, to call help center

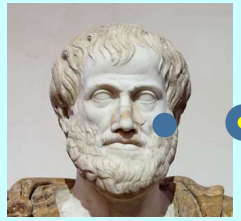
- **Does any AI/Robot system want to learn? (Does Alpha-Go want to play Go?)**

  NO: *None of the systems want to learn for themselves.*

  Simply, human programmers prepare them to learn

- **Does any system have self-consciousness?**

  NO: *None of the systems have self-consciousness*



Lifeform exists for soul

I think, therefore I am

- **If the goal is to build artificial lifeforms, we should aim to design artificial souls**
- **Then, is it a good idea to design artificial souls?**

  YES: *We need artificial souls for real partner robots*

Two types of artificial soul

Terminator type*:*

Enemy concept

- To compete and replace human
- Complete autonomous system
- Artificial intelligence

Doraemon type*:*

Friend concept

- To cooperate and help the human
- Augmenting human physical and intellectual ability
- Augmented intelligence

Robot Soul for a human friendly robot

- Robots should be friends of human
- Human is the master and the creator of robots
- Robot and human should cooperate with each other
- Robot is an embodiment of AI, but AI does not mean ***Artificial Intelligence***, but AI means ***Augmented Intelligence***

Reddy *personal communication 2017*

Summary

- New application areas, such as service robotics and partly field robotics, require dense contact between robots and users
- The robot is not necessary to be a complete autonomous system in those scenarios
- The robot should augment the intellectual and physical capability of human workers and or human users
- **Augmented Intelligence** instead of **Artificial Intelligence**
  - help human users with fellow human workers by augmenting human capabilities
  - Not to aim to replace human works but to aim lighten the burden of human workers
  - Ask the help to a fellow human worker, when encounters exception cases
  - Essential to be connected to a cloud system with cognitive service and overcoming the frame problem

View from the eastern culture

- Infinite loop of reductionism and holism
- Create new robotics and AI disciplines through the loop
- Robotics and AI in Trisahasramahasahasaralok adhatu (3000 big worlds)
- To create robot soul and to understand human soul