

Artificial Intelligence: Psychological & Societal Themes

Petros A. M. Gelepithis
(Πέτρος Α. Μ. Τζελεπίδης)

Plan of the talk

1. Hominisation.
2. Overview of Artificial Intelligence.
3. Societal theme: Decision making.
4. Psychological theme: Human Nature.
5. Epilogue.

Hominisation:

The Evolution of Human Species

Evolution of the artificial component of intelligence

(Gelepithis 2001, p. 28)

Appearance of Evolution
on planet Earth

Human Society at crossroads.

Appearance of Homo Habilis.
First use of stones as tool.
Probable appearance of design ability.

Appearance of **written** language.

Mechanisation of aspects of
Human knowledge.
Production of knowledge by AIS.

Appearance of Mathematics

??m YbN

2m YbN

000's YbN

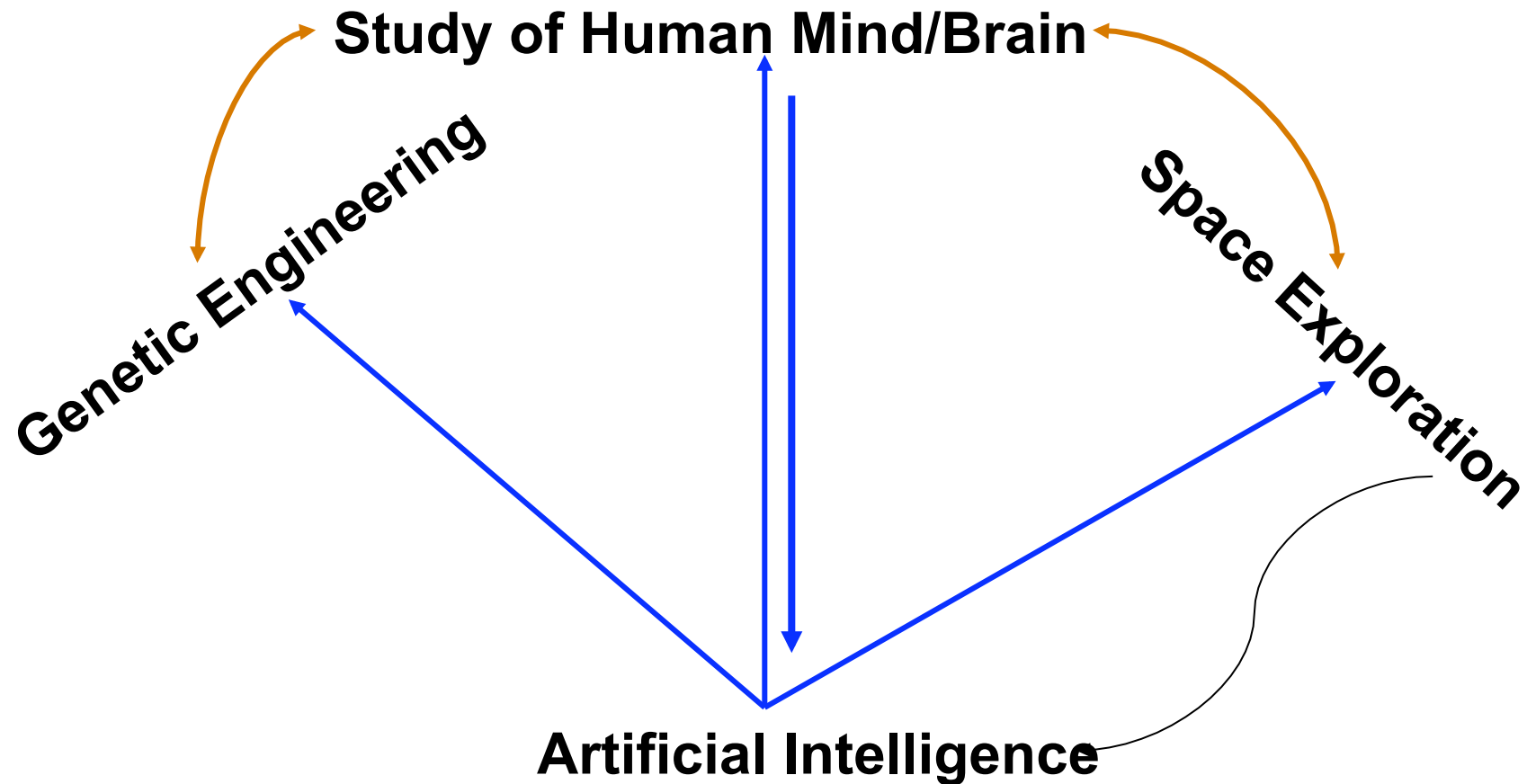
~2500 YbN

20th C.

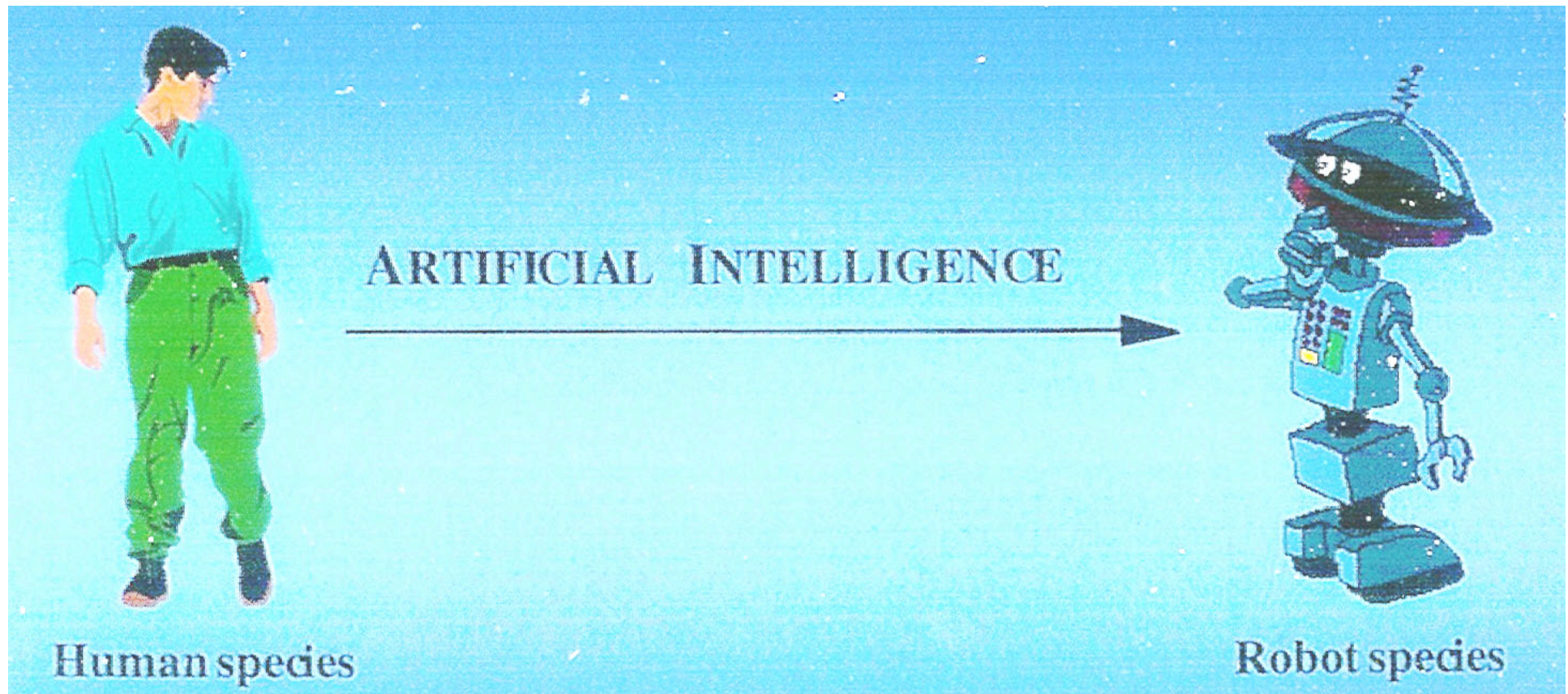
21st C.

YbN=Years before Now.

R&D Programmes of Planetary Significance



Όραμα, Εφιάλτης, ή Εξελικτική Αναγκαιότητα;



Plan of the talk

1. Hominisation.
2. Overview of Artificial Intelligence.
3. Societal theme: Decision making.
4. Psychological theme: Human Nature.
5. Epilogue.

Definition of Artificial Intelligence

The branch of human knowledge that deals with the study, design and development of **artificially intelligent systems** (AIS).

Kinds of AIS

Human-based
across a vast range of behaviour.

Formalisability
of Human Knowledge

Robot-based

**They do not exist.
But, they are feasible
-In principle.**

Examples of AIS
in
Science, Industry, Arts

AIS in Science

1. DENDRAL

2. AM

3. ID3

AIS in Science

4. Deep Blue

Defeated the world chess champion (May, 1997).

5. Palomar Observatory:

The AI Machine Learning subsystem succeeded in classifying heavenly bodies that were too faint to be classified by humans.

AIS in Industry

1. American Express

Credit decision making for debatable cases.

2. Westinghouse

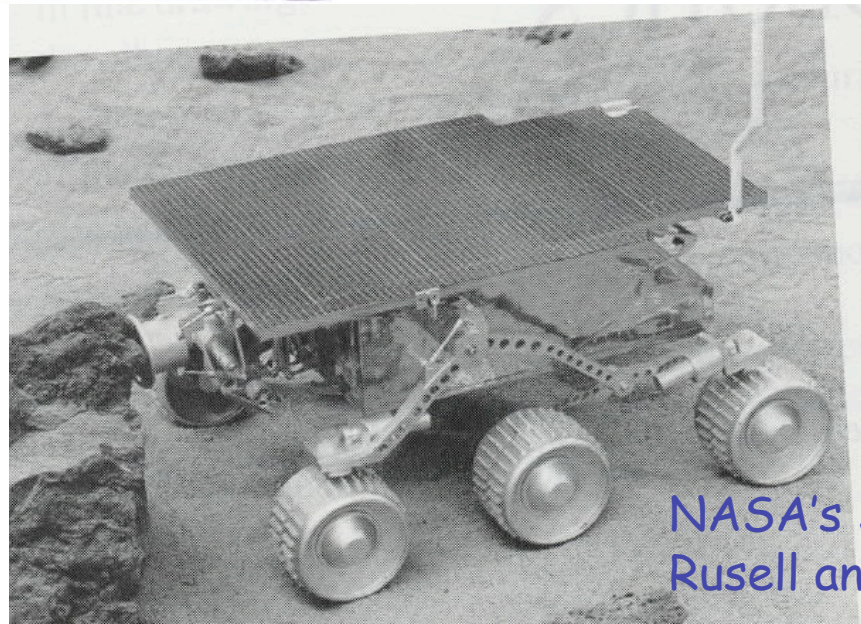
Expert system for fuel saving in nuclear stations.

AIS in Industry

3. Deputy Pilot (DARPA):
6 hierarchically composed Expert Systems.

4. Philips Speech Processing System:
Employed in more than 20 hospitals in Germany
and Austria until 1998.

5. Diamond cutting.



NASA's Sojourner.
Russell and Norvig, 2nd ed., 2003
achronic Gallery, Cyprus, 16/06/06

α νέα «παιδιά» του Ασίμοφ

KLAUS, by Volkswagen.

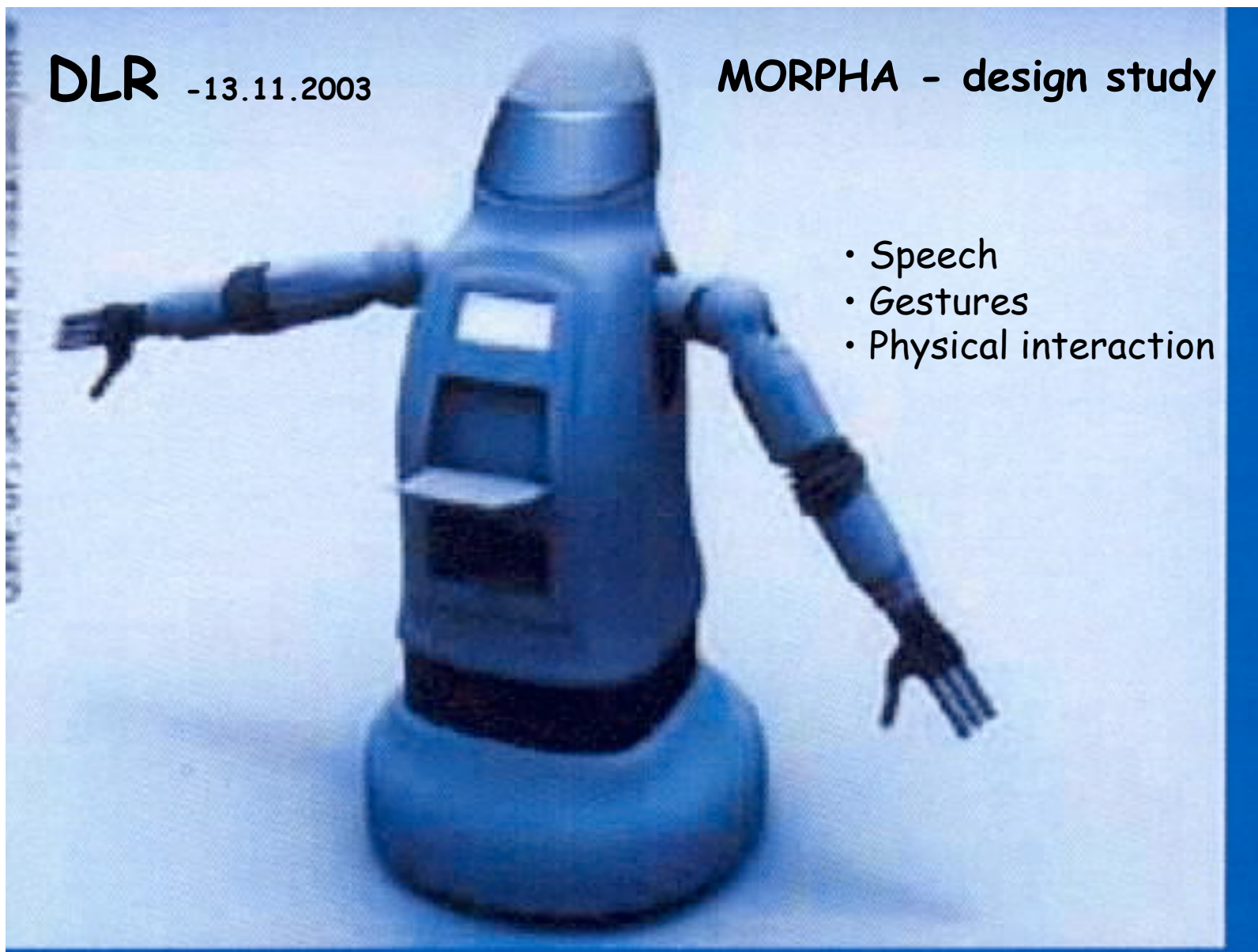
ΒΗΜΑΓΑΖΙΝΟ, 2002, π. 76



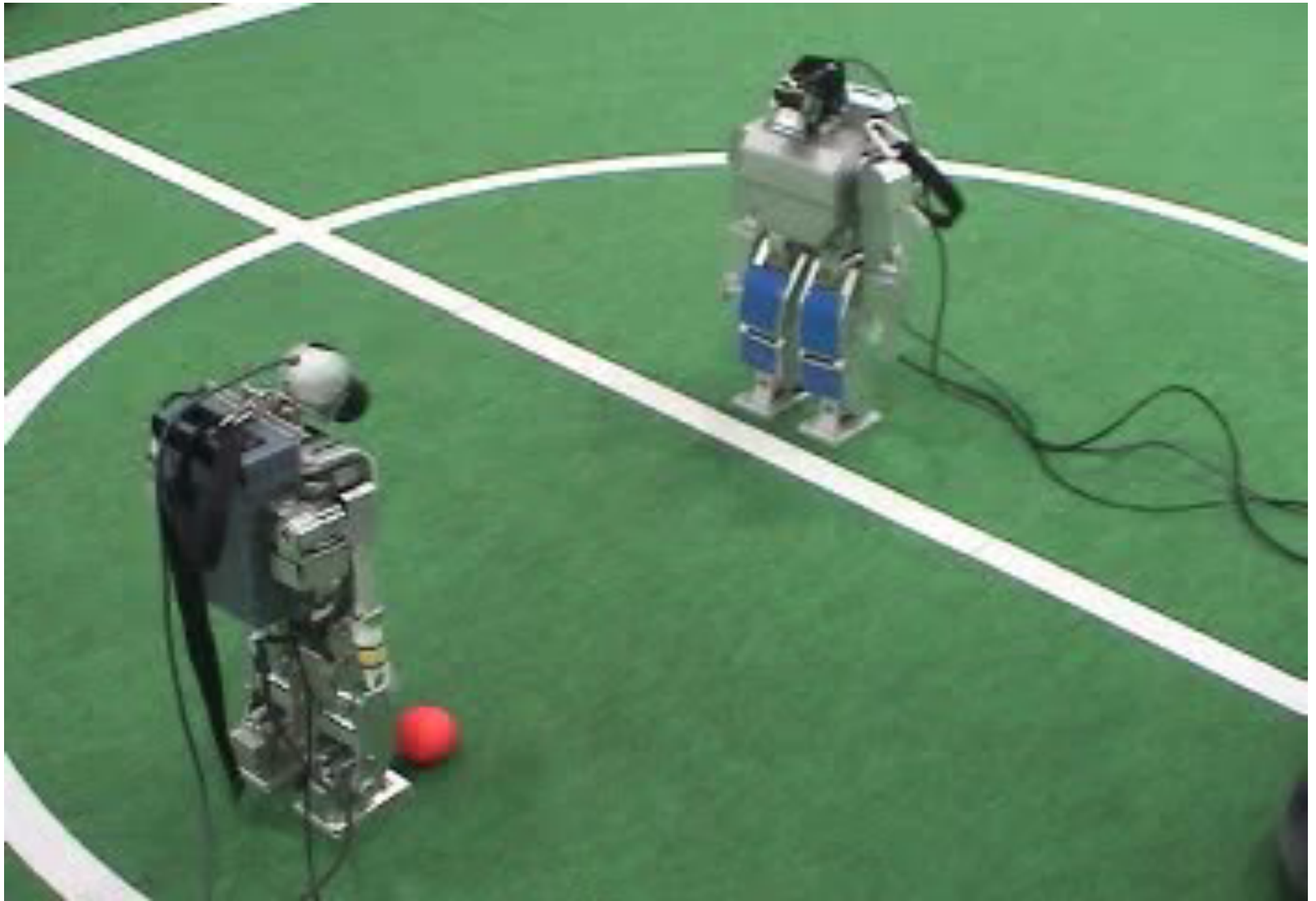
11

ALVIN/RALPH

Diachronic Gallery, Cyprus, 16/06/06







Diachronic Gallery, Cyprus, 16/06/06



ROBOVIE

Copyright ©

17



Copyright © PAMG, 1995-2005.

18

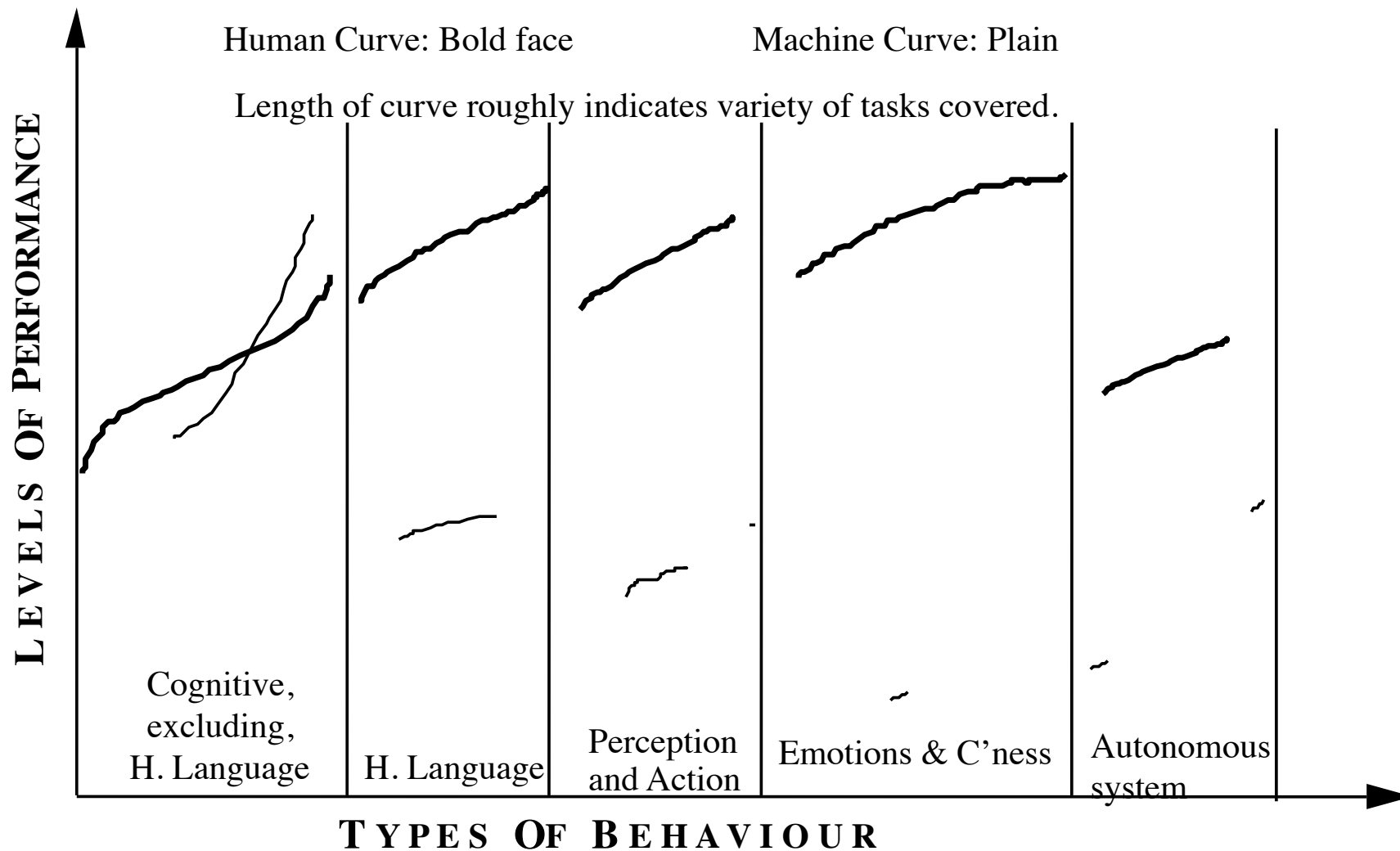
Diachronic Gallery, Cyprus, 16/06/06

AIS in Art

Cohen & AARON

Comparison of H. and Machine Intelligence in 1999

Gelepithis 1999, 2001.



Human Intelligence
is NOT
the highest intelligence.

AI: Payoffs and Ethics

“**in the Persian Gulf War**, one AI application used in targeting and scheduling of air flights was judged to be enormously valuable. The **DOD** [US Department of Defence] determined that **this single application** produced such significant savings, not to mention its effect on accelerating operations, that it **more than compensated for the total DOD investment, through all time**, in the field of AI.” (Hayes-Roth, 1997, p 109).

Exploitation of Basic Human Needs



So

us, 16/06/06

Plan of the talk

1. Hominisation.
2. Overview of Artificial Intelligence.
3. Societal theme: Decision making.
4. Psychological theme: Human Nature.
5. Epilogue.

The human-machine boundary

1. To what extent should our society incorporate **machine decision making**?
2. What criteria should we use to draw the *appropriate boundary* between human and machine responsibilities?

Question-1

Who should be responsible
whenever
an AIS makes a mistake?

Question-2

Making a mistake usually asks for an explanation. Would the machine-derived explanation be **comprehensible by**, and **sufficient to**, all those concerned?

Question-3

Explanations

depend on the premises of the system.

What types of knowledge
should be encoded in such premises?

Proposal

Since AIS are (and increasingly going to be) allowed to make considerable societal decisions, **then**, in a democratic society, citizens should be able to scrutinise **the premises** encoded in such systems.

Consequences

1. More transparent and reasonable government decision making.
2. Accountability of government officials easier and less evasive.

Fact

Learning machines can
be opaque with respect to their workings and
of limited predictability with respect to their
behaviour.

Remark

- a) Machine-created Knowledge may unduly influence human judgement.

- b) Machine-created Knowledge may produce individually or socially undesirable solutions to problems.

Question-4

Are we going to accept,
and eventually institutionalise,
the opaqueness and unpredictability of AIS?

Question-5

What would happen
when
human emotions and ethical values
are incorporated
in the premises of an AIS?

Question-6

Could we, in such a case,
distinguish between
Human and Machine decision making?

If different, which suggestion should we
follow?

Question-7

What should be the **scope of
Robot responsibilities?**

What should be the **scope of
Human responsibilities?!**

Proposed Decision making Criterion

Drawing the H-M boundary
should, primarily, be based on
'societally-useful' decision making.

Definitional Clarification

By 'societally-useful' decision making I mean that decisions concerning the well being of humans must incorporate the condition that no deterioration in the way of life of any human is allowed if that deterioration:

- (i) is not a result of wrong-doing by the individual concerned, or
- (ii) it has not been agreed by her; or
- (iii) it is not counterbalanced by a greater deterioration in the welfare of another individual or group.

Plan of the talk

1. Hominisation.
2. Overview of Artificial Intelligence.
3. Societal theme: Decision making.
4. Psychological theme: Human Nature.
5. Epilogue.

Potential impact of
the emerging machine intelligence on
Human Nature

Claim

Theories in
astronomy, biology, and psychology
have reduced, if not obliterated,
the significance of human beings.

Similarly

AIS will give human self-respect
the death blow

if humans at large perceive themselves as
nothing but machines.

Anti-theses

The realisation that human intelligence
is neither unique nor supreme
should have
a sobering effect on human arrogance.

Anti-theses

The advent of AI should be seen as
an ongoing breakthrough development
able to contribute greatly to
the well being of the human species.

Plan of the talk

1. Hominisation.
2. Overview of Artificial Intelligence.
3. Societal theme: Decision making.
4. Psychological theme: Human Nature.
5. Epilogue.

Hominisation: What's the next stage?

1. Extinction.
2. New species through natural evolutionary processes.
3. New species through genetic engineering.
4. New species through Artificial Intelligence.
5. **Hybrid development.**

Μωρέ, τι μηχανή είναι ο άνθρωπος!
Της βάζεις
ψωμί, κρασί, ψάρια, ραπανάκια,
και βγαίνουν
αναστεγμοί, γέλια και ονείρατα.

Νίκος Καζαντζάκης,
Βίος και Πολιτεία του Αλέξη Ζορμπά, σελ. 261.

Artificial Intelligence,
if successful, may well
immortalise the Human Spirit.