ROADS TO HUMAN LEVEL AI?

biological—imitate humans. Even neural nets, sho
work eventually.

engineering—solve problems the world presents— pres
ahead

direct programming, e.g. genetic algorithms

use logic, loftier objective. If it reaches hum
level, we will understand intelligence, and so will
robots. Logical AI has partly solved some inevitable pr
lems that haven't even been noticed by physiological
proaches.

The logic approach is the most awkward—except for
the others that have been tried.

Logic in AI

Features of the logic approach to AI—starting in 195

- Represent information by sentences in a logical guage, e.g. first order logic, second order logic, mo logic.

- Auxiliary information in tables, programs, states, is described by logical sentences.

- Inference is logical inference—deduction supplemen by calculation and some form of nonmonotonic infe 1980.

- Action takes place when the system infers tha
should do the action.

- Observation of the environment results in senter
in memory.

Topics, methods and problems of logical AI

- deduction, nonmonotonic reasoning, theories of
tion, problem solving,

- The frame, qualification, and ramification probl
have been partly solved.

- concepts as objects, contexts as objects, approxim
objects

- Elaboration tolerance (educate without brain surge

THE COMMON SENSE INFORMATIC SITUATIO

• Common sense: A structure composed of abilities knowledge. "Programs with common sense"—1959.

• The *common sense* informatic situation, which dif from a *bounded* informatic situation, has been difficul define precisely.

• Bounded informatic situations, e.g. chess positio take specific facts into account. In common sense, th is no limitation on what objects and facts may beco relevant.

• New facts may require revising conclusions, plans, algorithms. Formal *nonmonotonic reasoning*, e.g. cumscription and default logic, are important tools representing common sense reasoning in logic.

• Actions and other events often have only partly kn
able effects. Often not even probabillistic models
available.

COMMON SENSE INFORMATIC SITUATION—2

- Specific theories, e.g. scientific theories, are embed in common sense. Skills are also embedded in comm sense.
- Common sense physics: When two objects collide, th is a noise. An object pushed off a table will fall to floor.
- Common sense psychology: A person comes to dis someone whom he thinks killed his fellow countrymen
- The facts behind many human abilities are not o narily expressed in language but are often expressibl language or logic.
- Common sense abilities: Grasp object being touch Recognize a surface of an object—the knife. Fumble plastic surface.

• Common sense facts: $In(Rpocket, Knife, Now)$
$\wedge(\exists x)(Plastic\text{-}feel(x) \wedge Surface(x, Knife))$,

• Human-level common sense requires representing
up-to-now mental state as an object and reasoning ab
it.

EMBEDDING A SCIENTIFIC FACT IN SITUATIO CALCULUS

• Scientific theories are embedded in common sense, the formulas are embedded in natural language.
• Galileo's formula $d = \frac{1}{2}gt^2$ can be embedded in a sim common sense theory of situation calculus by

$$Falling(b, s) \wedge Velocity(b, s) = 0 \wedge Height(b, s) = h \wedge \wedge$$
$$d = \frac{1}{2}gt^2$$
$$\rightarrow (\exists s')(time(s') = time(s) + t \wedge Height(b, s') = Heigh$$

• For controlling a robot (1) must be used in connect with facts about concurrent events.
• The situation calculus formula connects Galileo's mula to quantities that are defined in (mostly unobserv situations to which the theory applies.

• Like other scientific formulas, Galileo's formula is u
more in constructing theories than in planning actio
a specific situation. Robots may use (1), i.e. $d = \frac{1}{2}$
expanded into situation calcculus, directly if they can
mediately measure the physical quantities involved.

EMBEDDING A SKILL IN COMMON SENSE—a philosophical path

- Objects exist independent of perception.

- Machine learning research is mistaken in concentrat on classifying perception. Herbert Simon's Bacon met for scientific discovery is limited by its concentration discovering relations among observables.

- A 3-d object is not a construct from 2-d views. learn about objects from view and by other means. T blind live in the same world as the sighted.

- Draw an object you can only feel but can't see. program that can get an object from a pocket is a g *Drosophila*.

EMBEDDING A SKILL IN SITUATION CALCULU

The skill of finding an object in a pocket can be pa
embedded in situation calculus.

$In(Knife, RPocket, s) \rightarrow Holding(Knife,$
$Result(Move(RHand, Interior(RPocket)); FumbleFor($
$Grasp(Knife); Remove(RHand, RPocket), s))$

Alternatively,

$In(Knife, RPocket, s) \rightarrow (\exists finger\ surface)$
$(surface \in Surfaces(Knife) \wedge finger \in Fingers(RHan$
$\wedge(\lambda(s')(Touches(finger, surface, s') \wedge Observes(Touche$
$surface), s')))(Result($
$Move(RHand, Interior(RPocket);$
$FumbleFor(Knife)), s)).$

Complications: Conscious guiding of the fumbling: fu ble until object is found, very little detailed informat is needed, and very little is available. For example, doesn't need information about the other objects in pocket.

Query: What do we know about the physics of pock and how is it represented in the human brain, and h should robots represent it?

"Keep trying a, and you will shortly achieve a situatio such that $Holds(f, s')$." How should this be represen logically?

SPECIFIC ABILITIES IMBEDDED IN COMMON SENSE

- Skills like walking, playing tennis
- Scientific theories
- AI programs, e.g. Mycin
- A chess player and a chess program
- Make a decision based determining which of two tions leads to a better resulting situation. Hum only maybe.

INTERACTION OF SKILLS AND KNOWLEDGE

- Partial knowledge of the skills and of situations.

- Picking my knife from my pocket containing coins keys.

- Interaction of observation with reasoning about act in logical AI, e.g. Filman, Reiter, Levesque, Sha han, Sandewall, Doherty.

- It may be new to emphasize partial knowledge ab effects of exercising a skill.

CURRENT PROJECT—DOMAIN DEPENDENT CONTROL FOR LOGICAL PROBLEM SOLVER

• General logical problem solvers without domain dep dent control experience combinatorial explosion.

• There is a profusion of cut-down logics.

• STRIPS should be a strategy for a logical prob solver. Likewise DASL.

• Minsky proposed in 1956 that a geometry theor prover should only try to prove sentences true in the agram. Herbert Gelernter implemented it, but in 1 IBM decided IBM should not be seen as making anyth but data-processing machines.

• Selene Makarios works on domain dependent cont She has some results in reducing search in the blo world.

EXAMPLES OF CONTROL

• When looking for feasible actions, don't subsitute i
formulas of the form $Result(a, s)$. This is part of STRI
• When trying to prove two triangles congruent and
have side a in one triangle equal to side a' in the ot
try to prove the corresponding adjacent angles equal.
• In blocks world and heuristically similar problems l
for moves to final position.
• Josefina Sierra-Ibañez and more recently Selene Ma
ios.

APPEARANCE AND REALITY

- The world is made of three-dimensional objects wh are only partly observable.
- History is only partly observable or even knowable.
- Reality is more persistent than appearance.
- Appearance of a scene after an event depends on reality of the scene and not just on what could be served.
- Pattern recognition and scientific discovery research not properly taken these facts into account.

ELABORATION TOLERANCE

• A collection of facts, e.g. a logical theory, is elaborat
tolerant to the extent that it can be readily elaborat
www.formal.stanford.edu/jmc/elaboration.html conta
theory and extensive examples.

• English language statements are very elaboration to
ant provided human common sense is available. Add
sentences will almost always work.

• Neural nets, connectionist systems, and present ch
programs have almost no elaboration tolerance. Exam
T. Sejnowski's Nettalk cannot be elaborated to incl
Pinyin pronunciations of the letters "x" and "q".

- Many elaborations of well constructed nonmonoto
logical theories can be accomplished just by adding s
tences.

- Example from "Missionaries and Cannibals". The
an oar on each bank of the river, ...

- *Formalizing Elaboration Tolerance* by Aarati Parma
a forthcoming Stanford dissertation.

FREE WILL IN A DETERMINIST WORLD

• We can make a situation calculus theory of a proc more determinist by adding axioms asserting that cert events occur.

• Human free will may consist of using a non-determi theory to decide deterministically on an action.

Here's a minimal example of using a non-determinist t ory within a determinist rule.

$$
\begin{aligned}
&Occurs(Does(Joe,\\
&\quad \textbf{if } \text{Better-for}(Joe, Result(Does(Joe, a1), s),\\
&\qquad\qquad Result(Does(Joe, a2), s))\\
&\quad \textbf{then } a1\\
&\quad \textbf{else } a2\\
&\qquad), s).
\end{aligned}
$$

• Here Better-for-Joe($s1, s2$) is to be understood as
serting that Joe thinks $s1$ is better for him than $s2$.

If we take the actor as understood, as is common in
uation calculus stdies, we get a shorter formula

$$\begin{aligned}
&Occurs(\textbf{if } \text{Better}(Result(a1, s), Result(a2), s)\\
&\qquad \textbf{then } a1\\
&\qquad \textbf{else } a2\\
&\qquad\quad), s).
\end{aligned}$$

• Do animals, even apes, make decisions based on co
paring anticipated consequences? If not, can apes
trained to do it? Chess programs do. According
Dan Dennett, some recent experiments suggest that a
sometimes consider the consequences of alternate
tions. Jane Goodall (personal communication) assu
me that chimpanzees do.

OTHER ASPECTS OF LOGICAL AI

- non-monotonic reasoning
- concepts as objects
- contexts as objects

REFERENCES

- Murray Shanahan, *The frame problem in artificial in-ligence* M.I.T. Press, 1997

- "What is artificial intelligence"–John McCarthy http:
formal.stanford.edu/jmc/whatisai.html.

- My AI articles are all on www-formal.stanford.edu/jm

- These references are completely inadequate.

- This research was supported by DARPA.