APPROXIMATE CONCEPTS AND APPROXIMAT THEORIES

John McCarthy
Computer Science Department
Stanford University
jmc@cs.stanford.edu
http://www-formal.stanford.edu/jmc/
March 28, 2006

My papers are on the above web page. This paper is http://www-formal.stanford.edu/approximate.html.

Our discussion will be adequate if it has as much clearn as the subject matter admits of, for precision is not be sought for alike in all discussions, any more than in the products of the crafts.—Aristotle

APPROACHES TO ARTIFICIAL INTELLIGENCE

biological—Humans are intelligent; imitate humans observe and imitate at either the psychological or neu physiological level

engineering—The world presents problems to intellige Study information and action available in the world.
1. Write programs that use non-logical representation
2. represent facts about the world in logic and dec what to do by logical inference

We aim at human level AI, and the key phenomeno the common sense informatic situation. Human leve is a difficult scientific problem, and many concepts needed. The concept of *approximate object* is one.

THE COMMON SENSE INFORMATIC SITUATIO

- Involves approximate entities, e.g. that don't h
  if-and-only-if definitions.

- There is no limitation on what information may
  relevant. Theories must be elaboration tolerant. E
  can easily add a condition that a particular boat ne
  a bailing can.

- Reasoning must often be non-monotonic. A too
  usable for its normal purpose.

SCIENCE IS EMBEDDED IN COMMON SENSE

Common sense theories contrast with formal scien
theories. However, science is embedded in common se

$$Falling(x,s) \wedge Velocity(x,s) = 0 \wedge \ d < Height(x,s)$$
$$\wedge\ d = \tfrac{1}{2}gt^2; \quad \text{Galileo's formula}$$
$$\rightarrow$$
$$(\exists s')(F(s,s') \wedge Height(x,s') = Height(x,s) - d).$$

Present AI theories are embedded in their designers' c
mon sense or in their users' common sense—not in a
system's common sense.

A LOGICAL ROAD TO HUMAN LEVEL AI

- Use *Drosophilas* that illustrate aspects of represen tion and reasoning problems.

- Concepts, context, circumscription, counterfactu consciousness, creativity, approximation

- narrative, projection, planning

- mental situation calculus

- domain dependent control of reasoning

EXAMPLES OF APPROXIMATE OBJECTS

- the rock constituting Mount Everest; climbed 1953
- taxable income—approximately defined according t

The Congress shall have power to lay and collect taxes incomes, from whatever source derived, without app tionment among the several States, and without reg to any census or enumeration.—16th Amendment

- $On(Block1, Block2)$,
- the wants of the United States,
- the present weight of John McCarthy,
- Theseus's ship with all new planks and a diesel eng
- the corner of a room,
- the boolean approximation of digital circuits.

APPROXIMATE OBJECTS AND APPROXIMATE THEORIES

Approximate objects

- don't have if-and-only-if definitions,
- usually don't have definite extensions,
- are prevalent in common sense,
- but not used in formal scientific theories,
- and not confessed in present AI theories,
- are needed for AI.
- Sentences about them can still have definite truth values.

THE US WISHED . . .

• The US wished Iraq would leave Kuwait. Not the sa
as George Bush wished.
• Very approximate useful entity.
• Bush statements $\rightarrow$ "US wished . . ."

• "US wished . . ." $\wedge$ Iraq in Kuwait $\rightarrow$ US action.
• The actual approximate objects are US wishes. A p
son's wishes are simpler.
• Ordinary fuzziness: weakly or strongly wished.

US WANTS IRAQ OUT OF KUWAIT

The following are all defaults.

$$\text{Says}(\text{entity}, \text{Wants}(\text{entity}, x)) \rightarrow \text{Wants}(\text{entity}, x),$$

$$\text{Says}(\text{President}(\text{USA}), x) \rightarrow \text{Says}(\text{USA}, x),$$

$$\text{Says}(\text{President}(\text{USA}), \text{Wants}(\text{USA}, \text{Leaves}(\text{Iraq}, \text{Kuwai}$$

$$\text{Wants}(x, y) \rightarrow (\exists z)(\text{Does}(x, z) \wedge \text{Achieves}(z, y)).$$

From these we infer

$$(\exists z)(\text{Does}(\text{USA}, z) \wedge \text{Achieves}(z, \text{Leaves}(\text{Iraq}, \text{Kuwait})$$

- This sketch leaves out a theory of action, distincti between objects and concepts, some necessary qualif tions, and depends on nonmonotonic reasoning. A theory includes these and a theory of $Wants$, $Says$, $Does$.
- It exemplifies the kind of inferences people make w approximate concepts.
- Computers will have to do something similar, becaus precise theory of "US wants" won't be available to co puters either.

APPEARANCE APPROXIMATES REALITY

• Reality is the situation s. Appearance is $app(s)$.

• The effect of an action a on reality is determined b
and s.

• The effect of an action a on the appearance $app(s$
not determined solely by a and $app(s)$.

• effect: $s' = result(a, s)$ noncommutativity
$\neg(\exists result1)(app(result(a, s)) = result1(a, app(s)))$

• The effect of action a on a part of appearance, nam
$app2(s) = subapp(\text{HORSES}, app(s))$ is determined b
and $app2(s)$

- limited commutativity:

$(\exists result2)(app2(result(a, s)) = result2(a, app2(s)))$

BLOCKS WORLD—sketch

- The real world with blocks and the finite, abstr
blocks world.
- Programs use the latter to decide what to do about
former.
- The abstract blocks world is an approximation to a c
text in the real world.
- The abstract representation needs to be *elaboration*
erant.
- A less approximate blocks world has one block in sc
position on another.

TWO THEORIES OF THE BLOCKS WORLD

The usual theory plus a theory in which one block on other has a displacement from being centered.

- Standard Theory: $Holds(On(b1, b2), s)$
- New Theory: $Holds(On(b1, b2, d), s)$

Relation between the theories:

- $Holds(On(b1, b2), s) \equiv (\exists d) Holds(On(b1, b2, d), s)$

PRELIMINARY REMARKS

- The philosophers' puzzle of Theseus's ship goes away if the ontology is rich enough to clude both the continuing entity and a specific col tion of planks.

- An approximate theory can become precise in a lim context—e.g. the blocks world.

- If it comes up, we'll argue or sue. Ambiguity of rules of *Go*.

PRELIMINARY CONCLUSIONS

- Solid logical towers on semantic quicksand.
- I don't yet see a deep theory of approximate obje
- For now, just predicate calculus with weak axiom
- Don't worry. Be happy.

WHAT TO DO WITH AN APPROXIMATE CONCE

- discover more—natural kinds, partial

- define more
- use as is
- add detail
- remove detail—maybe
- make a well-defined theory and relate to it

e.g. two blocks world theories, two travel theories

- improved numerical approximation
- better defined set boundary
- elaborated detail

QUESTIONS

- What kind of semantics is appropriate? Some appr
  imate theories may not admit semantics.

- When is an approximate concept an approximation
  something else?

- What kinds of approximations are there?

- When is an approximate theory good enough for u

- How should an AI system think about the fact of
  proximation?

- What are the relations between approximations?

WHEN IS AN APPROXIMATE THEORY GOOD ENOUGH FOR USE?

- Observations of the world, e.g. with a TV camera by a robot putting its hand in your pocket) give to sentences and sometimes other data structure the memory of the robot. The sentences are appr imations to reality.

- Sometimes the robot reasons to *understand* the s ation. The conclusions are approximate.

- The robot decides what to do. The action comma are schematic.

- The action hardware, e.g. an arm takes an act
  action in the real world. Depending on the quality
  the theory and the real situation, the action may
  appropriate.

- When a complication occurs in the real world
  proves important, the theory must be elaborated, p
  ably in real time.

CLAIMS

- If-and-only-if definitions are rare.
- The common sense informatic situation involves appr
imate concepts.
- Formalized scientific theories rarely do.
- Numerical approximation is rare.
- Fuzziness is one kind of approximation.
- Human reasoning only partly corresponds to mathem
ical reasoning.
- **But mathematical logic is the right tool.**
- We use non-monotonic reasoning, approximate obje
contexts as objects.
- Natural kinds are relative to context.

HARANGUE

- Intelligence is the ability to understand complica environments and decide what to do.

- The basic problems are similar for humans, robots Martians.

- The scientific goal of AI research should be hur level intelligence, not immediate shrink-wrapped plications.

- Computer science has been somewhat corrupted the dot com frenzy for demos to attract investors to compete for Government money.

- The same problems arise in logical AI and biol
  cal AI, but they have been better understood by
  people in logical AI. The frame problem, the qua
  cation problem and the ramification problem are
  vious examples. Present biology motivated syste
  have almost no elaboration tolerance, e.g. Mycin
  NetTalk. Most of the biology oriented AI research
  don't even recognize the problems.

- AI researchers should consider the approximate nat
  of the concepts used in the formalsm and its relat
  to reality and less approximate concepts.