

Some Philosophical Problems From The Standpoint Of Artificial Intelligence

John McCarthy and Patrick J. Hayes

Presented by Xiaodong Yu

What is Artificial Intelligence?

An intelligent entity that is equipped with a representation or model of the world. On the basis of this representation, a certain class of internally posed questions can be answered. For example:

- What will happen next in a certain aspect of the situation?
- What will happen if I do a certain action?
- What is $3 + 3$?
- What does he want?
- Can I figure out how to do this or must I get information from someone else or something else?

Adequacy for Representations of the World

- **Metaphysically adequate:** if the world could have that form without contradicting the facts of the aspect of reality that interests us. For example:
 - The representation of the world as a collection of particles interacting through forces between each pair of particles.
- **Epistemologically Adequate:** it can be used practically to express the facts that a person actually has about the aspect of the world. For example: “John is at home”, “John's telephone number is 321-7580”.
- **Heuristically Adequate:** if the reasoning processes actually gone through in solving a problem are expressible in the language.

Formalism

Formal notion of an epistemologically adequate system:

- Situations
- Fluent
- Causality
- Actions
- Strategies
- Knowledge and Ability

Situations

- Situation is the complete state of the universe at an instant of time
- Universe is too large. Impossible to describe completely. But can determine some of the facts about the situations.
- For example, “a certain person has offered \$250 to buy Mr. Smith’s car”.
- We only need to know that \$250 is too low for a car.
- We only need partial information about a situation.

Fluents

- A fluent is a function whose domain is the space of situations
- Propositional Fluent: the range of fluent is true or false.
- Situational Fluent: the range is the space of situations.
- For example: assert about a situation s that person p is in place x and that it is raining in place x
- $[\lambda s'.at(p, x, s') \wedge raining(x, s')](s)$

Causality

- a fluent $F(\pi)$ where π itself is a propositional fluent.
- $F(\pi, s)$ asserts that the situation s will be followed by a situation that satisfies the fluent π .
- For example, to assert if a person is out in the rain he will get wet
- $\forall x.\forall p.\forall s.raining(x, s) \wedge at(p, x, s) \wedge outside(p, s) \rightarrow F(\lambda s'.wet(p, s'), s)$

Actions

- Actions is a situational fluent $\text{result}(p, \sigma, s)$
- p : person; σ : an action; s : a situation
- The value of $\text{result}(p, \sigma, s)$ is the situation that results when p carries out σ , starting in the situation s .
- For example, to assert if a person with a key that fits the safe can open the safe.
- $\text{has}(p, k, s) \wedge \text{fits}(k, sf) \wedge \text{at}(p, sf, s) \rightarrow \text{open}(sf, \text{result}(p, \text{opens}(sf, k), s))$

Strategies

- Actions can be combined into strategies.
- Consider for example the strategy that consists of walking 17 blocks south, turning right and then walking till you come to Chestnut Street.

- **begin**

s := *result*(*p*, *face*(*South*), *s*);

n := 0;

b : **if** *n* = 17 **then go to** *a*;

s := *result*(*p*, *walk-a-block*, *s*);

n := *n* + 1;

go to *b*;

a : *s* := *result*(*p*, *turn-right*, *s*);

c : *s* := *result*(*p*, *walk-a-block*, *s*);

if *name-on-street-sign* ≠ '*Chestnut Street*' **then go to** *c*

end;

Knowledge and Ability

- To achieve some goals, we need some knowledge.
- Go back to the safe example.
- $has(p, k, s) \wedge fits(k, sf) \wedge at(p, sf, s) \rightarrow open(sf, result(p, opens(sf, k), s))$
- Now the safe become a combination safe, and we need a combination to open it instead of a key. The problem becomes
- $at(p, sf, s) \wedge csafe(sf) \rightarrow open(sf, result(p, opens2(sf, combination(sf)), s)).$
- P might not know the combination.
- $open(sf, result(p, opens2(sf, idea-of-combination(p, sf, s)), s))$

Far from Epistemological Adequacy

- The approximate character of $\text{result}(p, \sigma, s)$: an action must lead to a definite situation, which is not true. For example:
- if someone is asked, 'How would you feel tonight if you challenged him to a duel tomorrow morning and he accepted?' he might well reply, 'I can't imagine the mental state in which I would do it; if the words inexplicably popped out of my mouth as though my voice were under someone else's control that would be one thing; if you gave me a long-lasting belligerence drug that would be another.'
- It is not clear how to take uncertainty into account by attaching probabilities to its sentences.

Far from Epistemological Adequacy

- Parallel Processing: many processes are going on simultaneously, so we need a formalism that allows parallel programs.

Conclusions

- A construction of formal notion of an epistemologically adequate system, but still far from epistemological adequacy.
- Because of the limit of time, More details are in the paper.

Thank you!