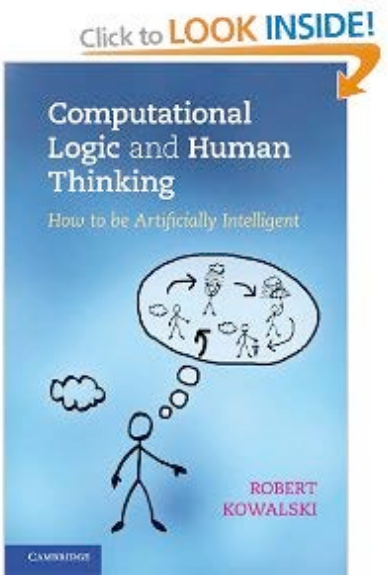
# Artificial Intelligence and Human Thinking

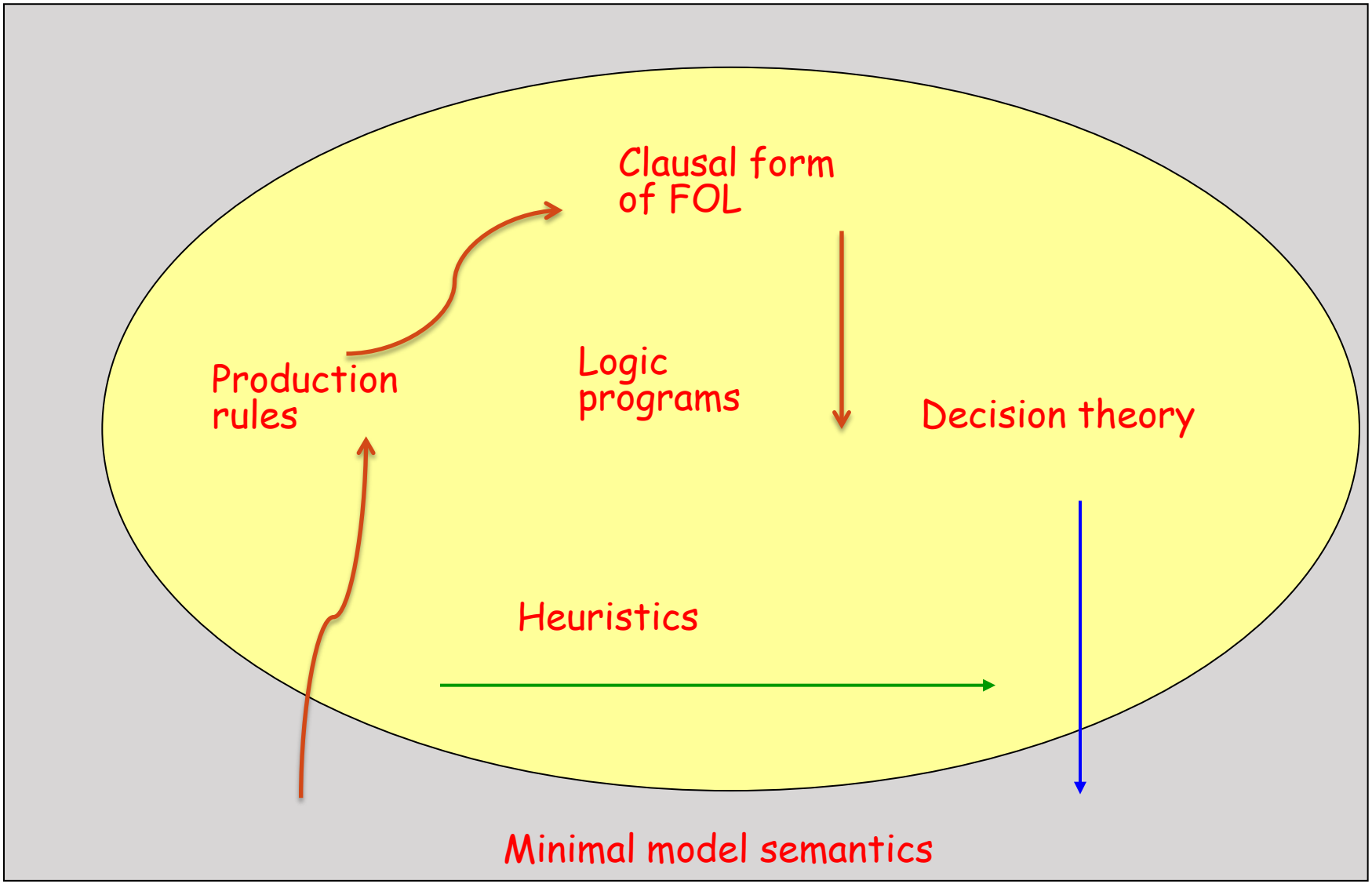
Robert Kowalski
Imperial College London



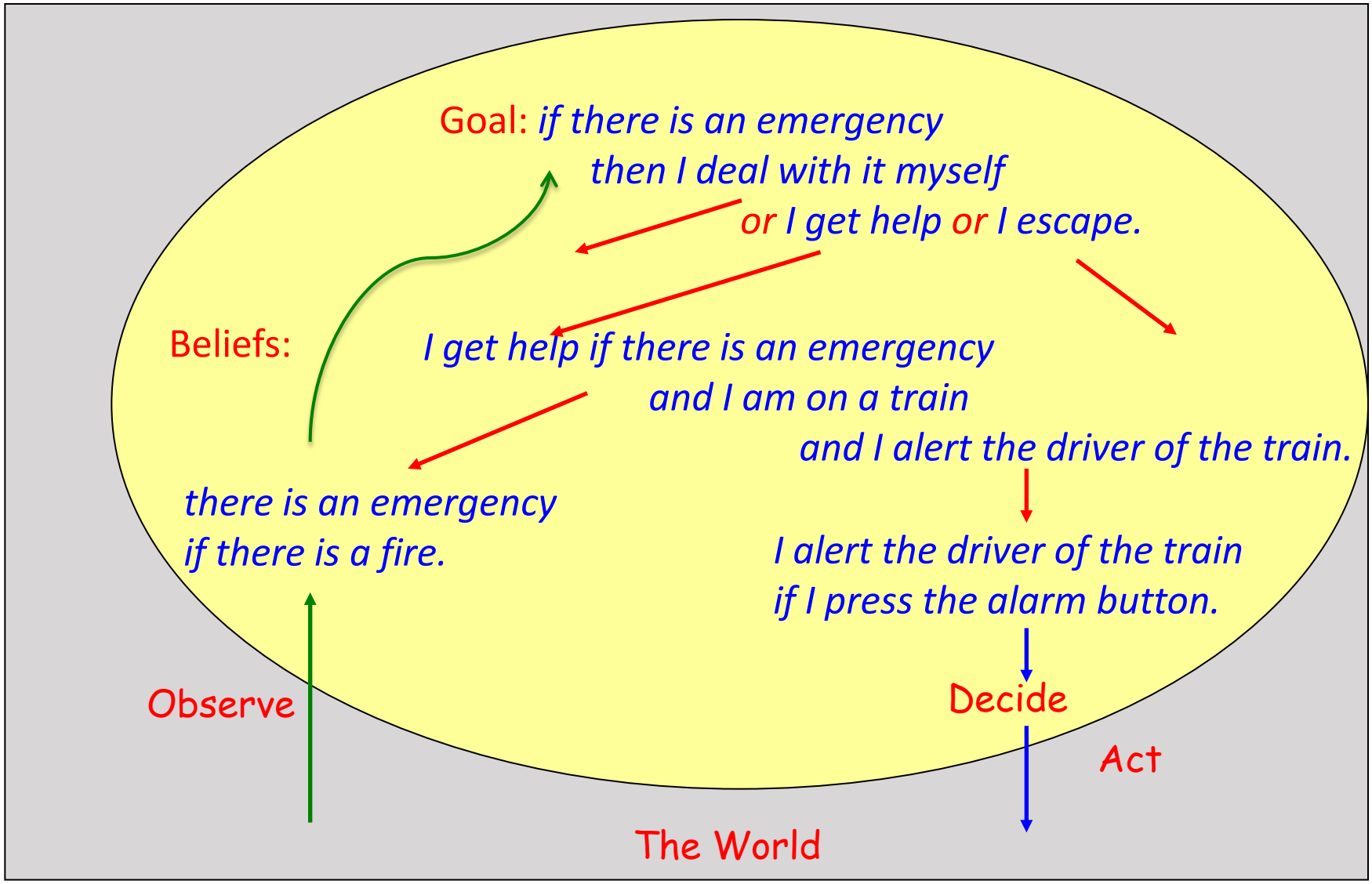
# Artificial Intelligence and Human Thinking

- The Abductive Logic Programming (ALP) agent model as a unifying framework
- ALP clausal logic as the Language of Thought (LOT)
- ALP clausal logic as a connectionist model of the mind

# The ALP Agent Model as a unifying framework



# An Agent on the London Underground



## Complex thinking and decision-making can be compiled into more efficient heuristics

For example: *if there is a fire*
*and I am on a train*
*and I can not deal with the fire myself*
*then I press the alarm button*.

## Lower-level heuristics and higher-level thinking and deciding can be combined, as in dual process models of human thinking.

As Kahneman and Frederick (2002) put it:

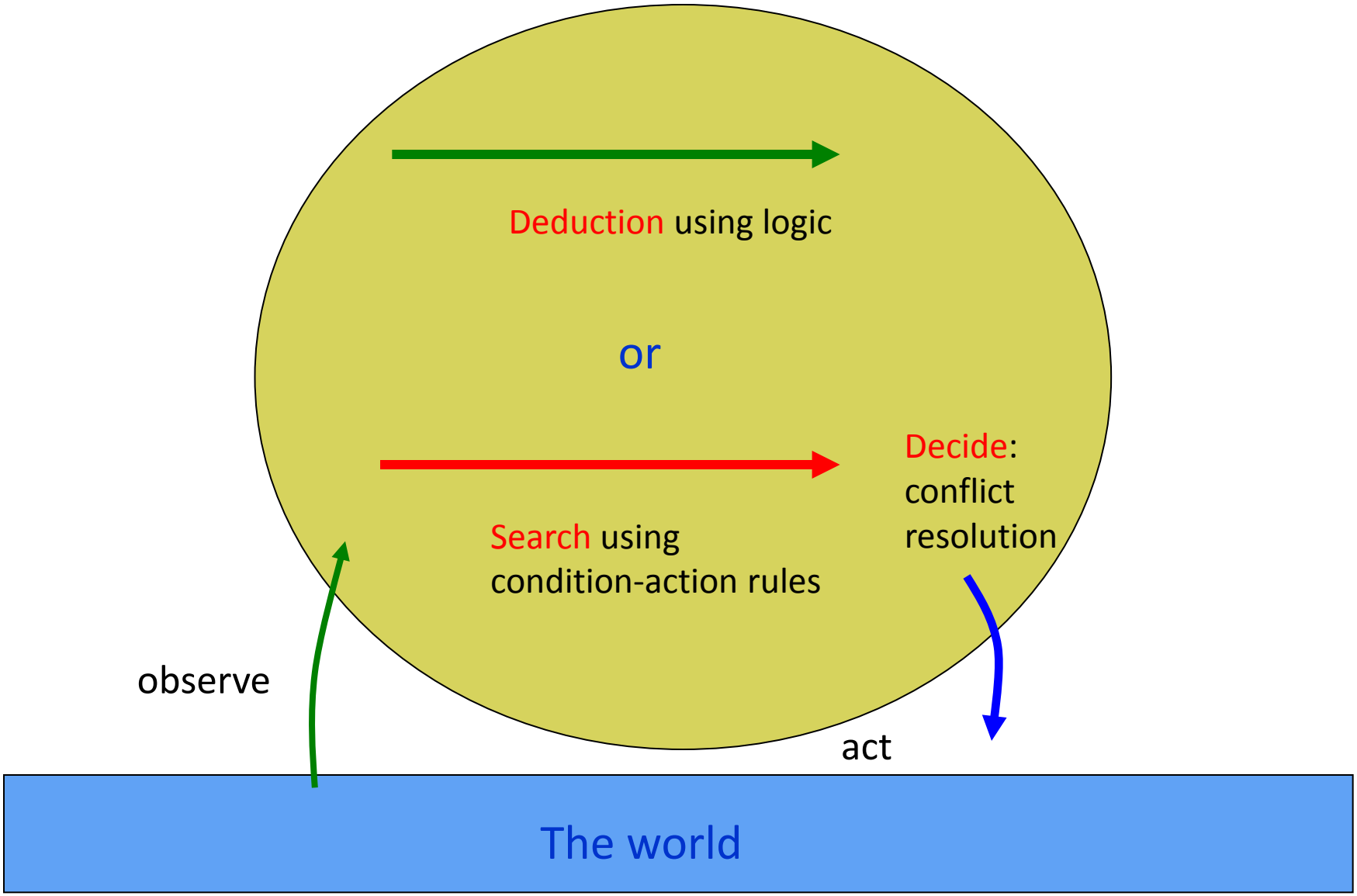
> the intuitive, subconscious level "quickly proposes intuitive answers to judgement problems as they arise",
> while the deliberative, conscious level "monitors the quality of these proposals, which it may endorse, correct, or override".

Paul Thagard in "Mind: Introduction to Cognitive Science" (1996) writes, page 45:

"In logic-based systems the fundamental operation of thinking is *logical deduction*, but from the perspective of rule-based systems the fundamental operation of thinking is *search*."

Here *rule-based system* means condition-action rules (or production rules) in production systems.

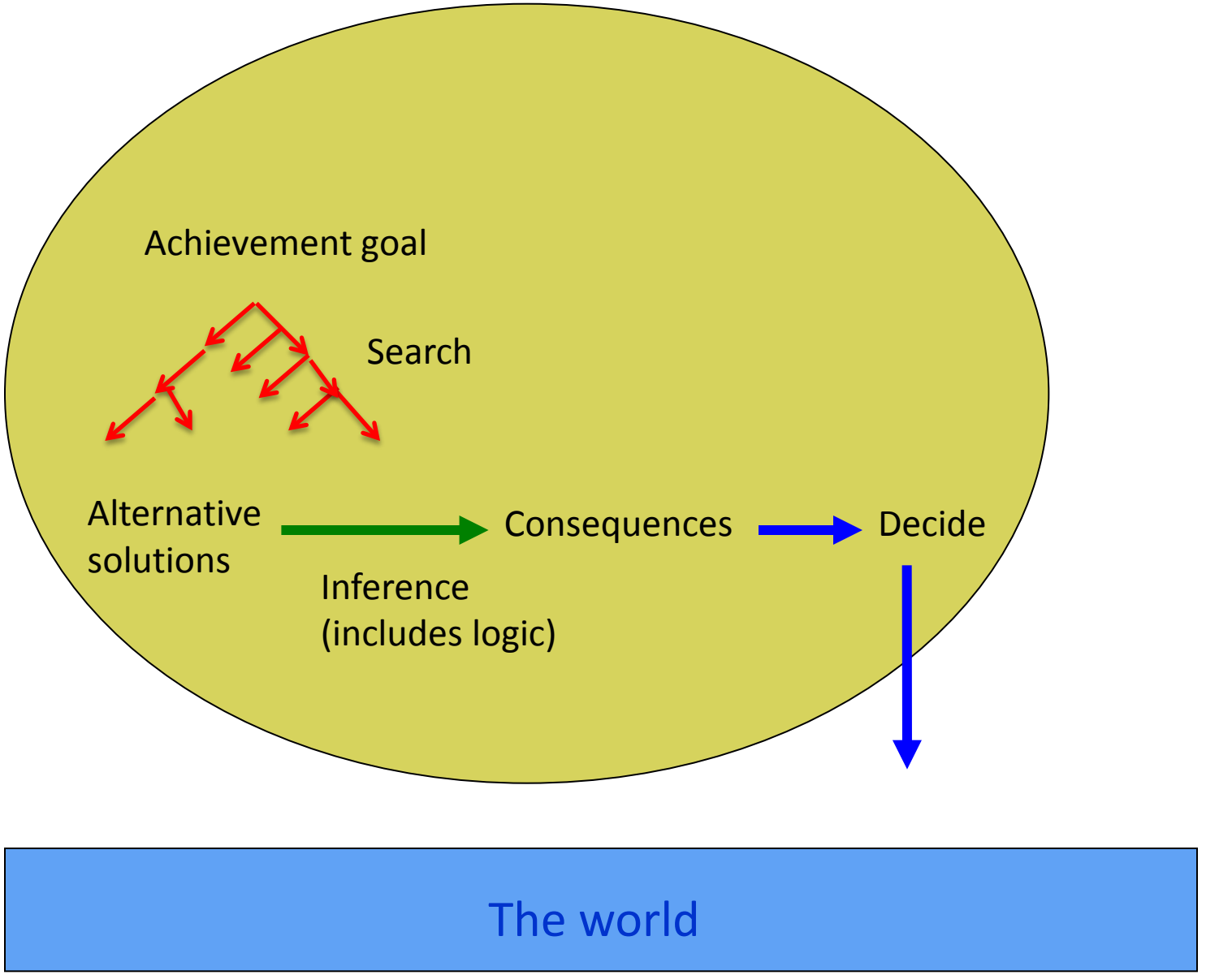
# Thagard's view of the difference between logic and production systems



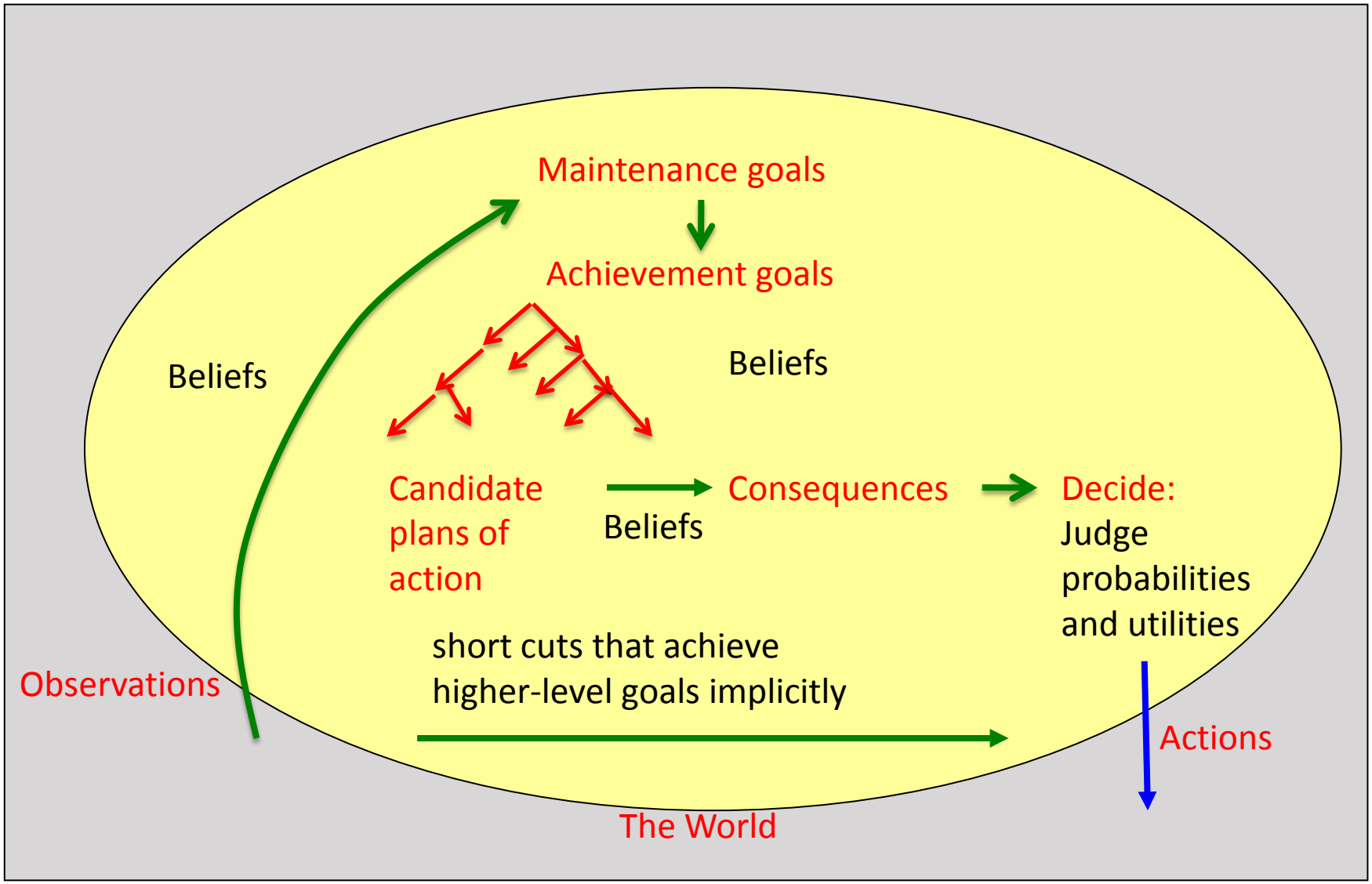
Jonathan Baron "Thinking and Deciding" (Fourth edition, 2008):

"*Thinking* about actions, beliefs and personal goals can all be described in terms of a common framework, which asserts that thinking consists of *search* and *inference*. We *search* for certain objects and then *make inferences* from and about the objects we have found." (page 6)

# Baron's view of search in relation to thinking and deciding



# Simplified Abductive Logic Programming (ALP) agent cycle



# The syntax of goals and beliefs in ALP agents

Beliefs are logic programs:

*conclusion if condition$_1$ …. and condition$_n$*

equivalently: *if condition$_1$ …. and condition$_n$ then conclusion*

Goals are clauses in first-order logic (FOL):

*If condition$_1$ …. and condition$_n$*
*then conclusion$_1$ …. or conclusion$_m$*

All variables are universally quantified.

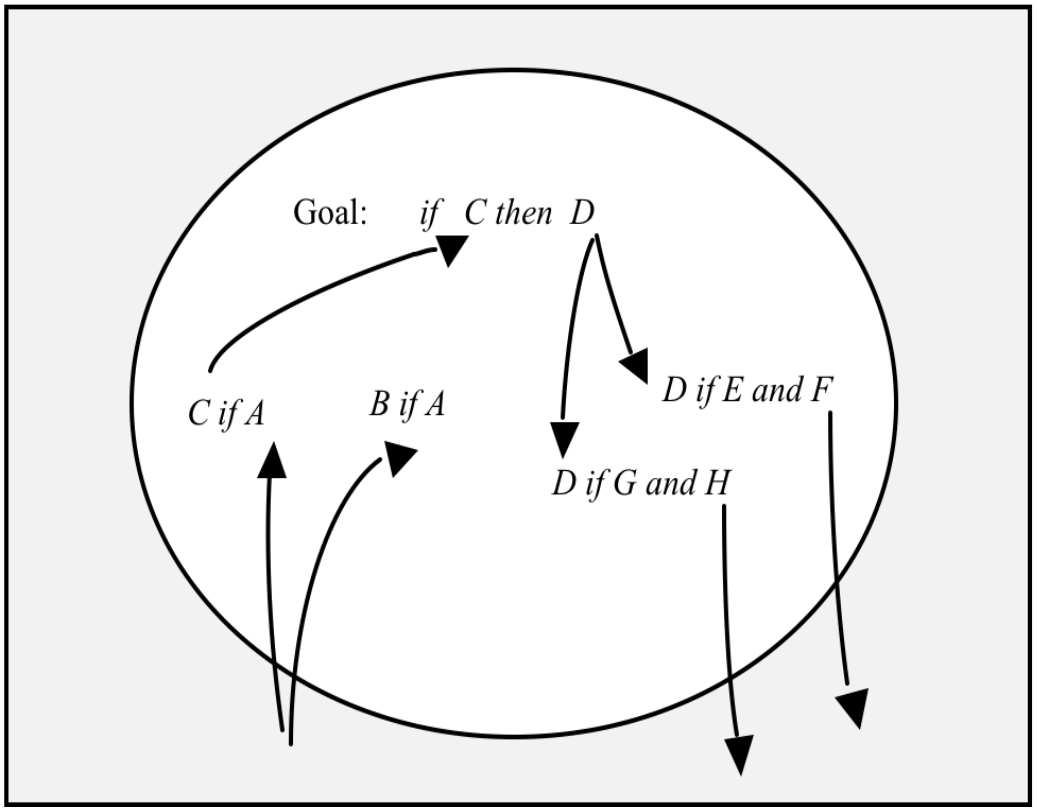
# ALP agents – minimal model semantics

Beliefs *B* describe the world as the agent imagines it.
Goals *G* describe the world as the agent would like it to be.

Given observations *O,*
the agent's task is to generate a set $\Delta$
of actions and assumptions such that:

$G \cup O$ is *true* in the minimal model determined by $B \cup \Delta$.

# Internal clauses and links need not represent states of affairs in the real world

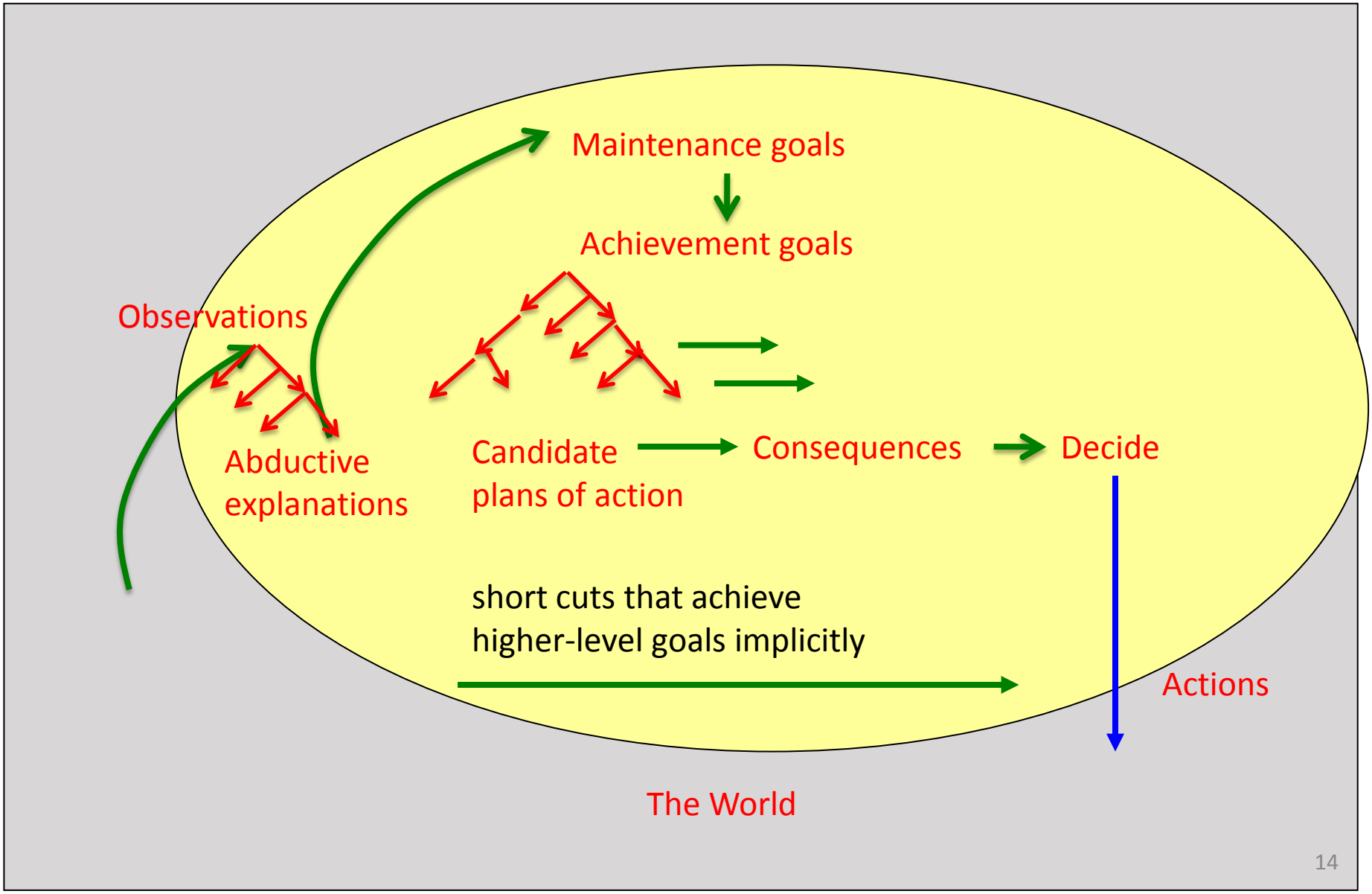



Only A, F and H correspond directly to reality.

C and D are mental constructs, which help the agent organise its thoughts.

The definition of truth *projects* these mental constructs onto the external world.

# Abductive Logic Programming (ALP) includes the abductive explanation of observations



Goal *G*: *if there is an emergency*
*then I deal with it myself*
*or I get help or I escape*

Observation *O*: *there is smoke*

Beliefs B: *there is smoke if there is a fire*
*there is an emergency if there is a fire*
*I get help if I press the alarm button*

$G \cup O$ is true in the minimal model determined by $B \cup \Delta$, where

$\Delta$= {*there is a fire, I press the alarm button*}.

explains *there is smoke*
(abduction)

achieves *I get help*
(planning)

# ALP can deal with uncertainty by associating probability with conditions

In general:

*A consequence takes place*
*if an action is performed*
*and the world is in a particular state*

For example:

*You will be rich*
*if you buy a lottery ticket*
*and your number is chosen* probability = .0001

*It will rain*
*if you do a rain dance*
*and the gods are pleased* probability = ?

# Different $\Delta$ can solve the same task.

The challenge is to find the best $\Delta$
within the computational resources available.

In classical decision theory, actions are evaluated
by measuring the expected utility of their consequences.

In philosophy of science, explanation are evaluated
by measuring their probability and explanatory power.
(The more observations explained the better.)

In ALP agents, actions and assumptions are combined in $\Delta$,
and are treated in the same way, and
forward reasoning is used to generate their consequences,

## The ALP agent model can help agents make better decisions.

In classical decision theory, all alternative actions and their consequences (outcomes) are assumed given in advance.

In ALP agents, alternative actions and their consequences are generated by using beliefs to search for solutions of goals.

The same evaluation criteria can be used both

to decide between alternatives and

to guide the search.

A theoretical framework for goal-based choice and for prescriptive analysis. Kurt A. Carlson & Chris Janiszewski & Ralph L. Keeney & David H. Krantz & Howard C. Kunreuther & Mary Frances Luce & J. Edward Russo & Stijn M. J. van Osselaer & Detlof von Winterfeldt. Market Lett (2008) 19:241–254.

"We view consumer preferences and consumer decisions as the output of goal pursuit. This departure from rational economic models allows us to characterize consumer behavior more fully.

For example, instead of assuming that consumer choices are the simple output of application of one's utility function to a set of known alternatives, with known consequences, we assume that choices result from consumers' attempt to satisfy numerous goals of different types (e.g., process, outcome, social), under a variety of constraints."

# Artificial Intelligence and Human Thinking

- The Abductive Logic Programming (ALP) agent model as a unifying framework
- ALP clausal logic as the Language of Thought (LOT)
- ALP clausal logic as a connectionist model of the mind

# ALP as the Language of Thought (LOT)

In the philosophy of language, there are three schools of thought :

The LOT is a private, language-like representation,
which is independent of public, natural languages.

The LOT is a form of public, natural language;
and the natural languages that we speak influence the way we think.

Human thinking does not have a language-like structure at all.

In ALP agents, clausal logic serves as an agent's private LOT,
independent of any public language.

# ALP clausal logic as the LOT

According to relevance theory [Sperber and Wilson, 1986],
people understand natural language by attempting
to extract the most information for the least processing cost.

It follows that:

If you want to identify the nature of the LOT,
then you should study communications that are easy to understand.

If you want your communications to be easy to understand,
then you should express them in a form that is close to the LOT.

# The Emergency Notice on the London underground

Press the alarm signal button to alert the driver.

The driver will stop
if any part of the train is in a station.

If not, the train will continue to the next station,
where help can more easily be given.

There is a 50 pound penalty for improper use.

# The Meaning of the London Underground Notice

*the driver is alerted*
*if you press the alarm signal button.*

*the driver will stop the train in a station*
*if the driver is alerted*
*and any part of the train is in the station.*

*the driver will stop the train in the next station*
*if the driver is alerted*
*and not any part of the train is in a station.*

*help can more easily be given in an emergency*
*if the train is in a station.*

*You may be liable to a £50 penalty*
*if you use the alarm signal button improperly*

ELIZABETH II

# British Nationality Act 1981

**1981 CHAPTER 61**

An Act to make fresh provision about citizenship and nationality, and to amend the Immigration Act 1971 as regards the right of abode in the United Kingdom.

[30th October 1981]

BE IT ENACTED by the Queen's most Excellent Majesty, by and with the advice and consent of the Lords Spiritual and Temporal, and Commons, in this present Parliament assembled, and by the authority of the same, as follows:—

## PART I

### BRITISH CITIZENSHIP

*Acquisition after commencement*

Acquisition by birth or adoption.

**1.**—(1) A person born in the United Kingdom after commencement shall be a British citizen if at the time of the birth his father or mother is—

(*a*) a British citizen ; or

(*b*) settled in the United Kingdom.

(2) A new-born infant who, after commencement, is found abandoned in the United Kingdom shall, unless the contrary is shown, be deemed for the purposes of subsection (1)—

(*a*) to have been born in the United Kingdom after commencement ; and

(*b*) to have been born to a parent who at the time of the birth was a British citizen or settled in the United Kingdom.

1.-(1) A person born in the United Kingdom after commencement shall be a British citizen if at the time of the birth his father or mother is –

(a) a British citizen; or

(b) settled in the United Kingdom.

The logic of subsection 1.-(1)

*A person shall be a British citizen by 1.-(1)*

*if* *the person was born in the United Kingdom*

*and* *the person was born after commencement*

*and* *a parent of the person was a British citizen at the time of the person's birth or a parent of the person was settled in the United Kingdom at the time of the person's birth.*

# The problems of understanding natural language communications

1. Identify the intended meaning of ambiguous sentences. e.g.

   he gave her the book.

   } clarity

2. Represent the intended meaning, e.g.

   John gave Mary the book.
   John gave the book to Mary.
   Mary received the book from John.
   The book was given to Mary by John.

   in a simple canonical form. e.g.

   } simplicity

   *give(john, book, mary, e1000)*

3. Connect the canonical representation
   with other mental representations,
   in a way that makes it easy to use the representation later.

   } coherence

# Clausal logic is a canonical form of FOL.

In clausal logic, sentences have a simplified form, e.g.:

*has-feathers(X) ← bird(X).*
*bird(john).*

In standard FOL, the same beliefs can be expressed in infinitely many, equivalent ways, including:

*¬(∃X((¬has-feathers (X) ∧ bird(X)) ∨ ¬bird(john)))*
*¬(∃X((¬has-feathers (X) ∨ ¬bird(john)) ∧ (bird(X) ∨ ¬bird(john))))*

In clausal logic, reasoning is simpler than in standard FOL and can be reduced to forward or backward reasoning, which are special cases of the resolution rule.

# Clausal logic as a model of the LOT can help people to communicate more effectively

By expressing communications:

| | |
|---|---|
| Clearly | So that their meaning is unambiguous. |
| Simply | So that their meaning is close to their canonical form. |
| Coherently | So that it is easy to link new information to old information. |

Joseph M. Williams

# Style

## Toward Clarity and Grace

*With two chapters coauthored by*
Gregory G. Colomb

The University of Chicago Press
*Chicago and London*

Contents

# Williams: Two Principles of Coherence

"1. Put at the beginning of a sentence those ideas that you have already mentioned, referred to, or implied, or concepts that you can reasonable assume your reader is already familiar with, and will readily recognise."

2. Put at the end of your sentence the newest, the most surprising, the most significant information: information that you want to stress – perhaps the information that you will expand on in your next sentence."

# Coherence

Example:

A.
If A then B.
If B then C.
Therefore C.

Example:

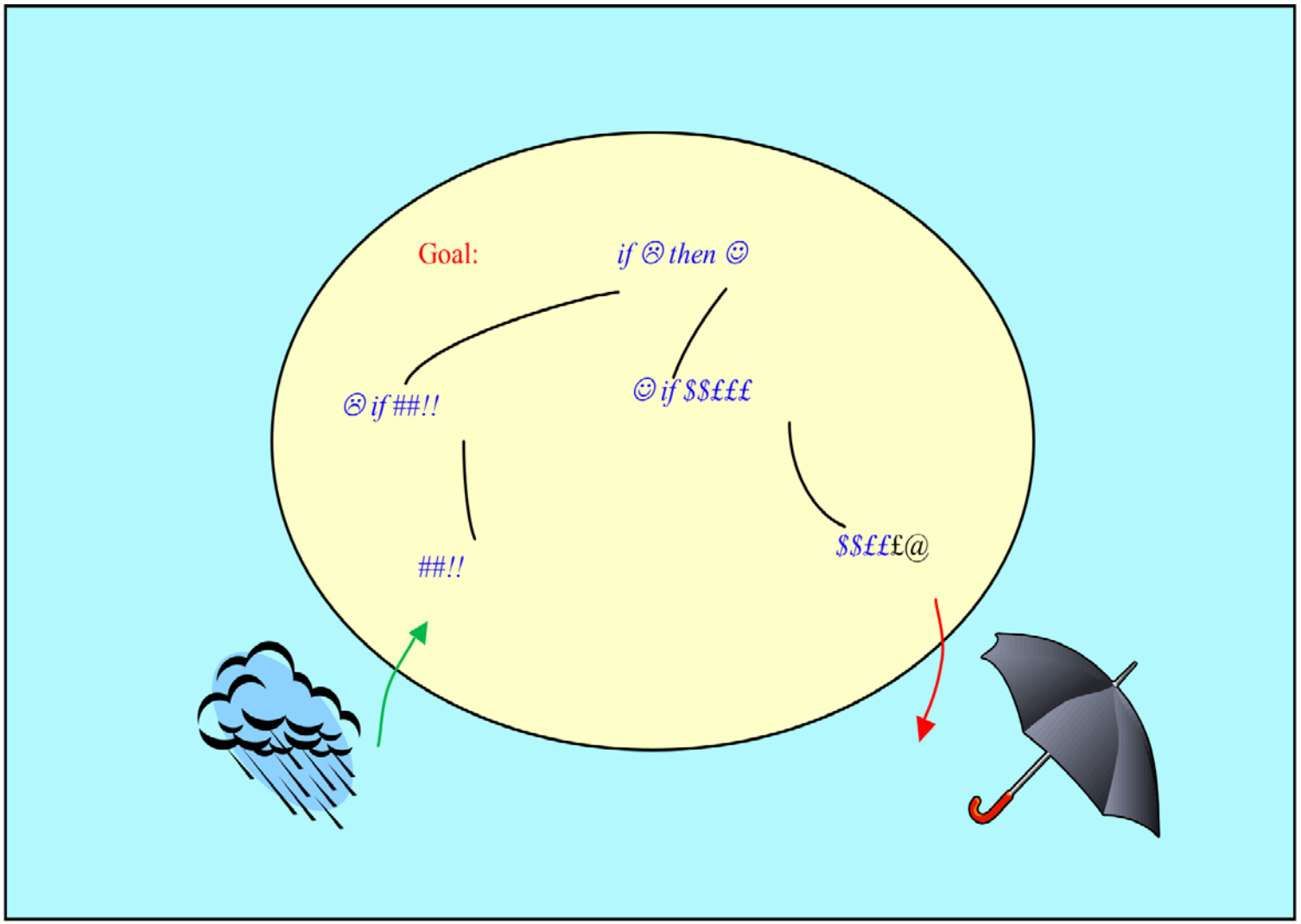
C?
C if B.
B if A.
A.
Therefore C.

# Artificial Intelligence and Human Thinking

- The Abductive Logic Programming (ALP) agent model as a unifying framework
- ALP clausal logic as the Language of Thought (LOT)
- ALP clausal logic as a connectionist model of the mind

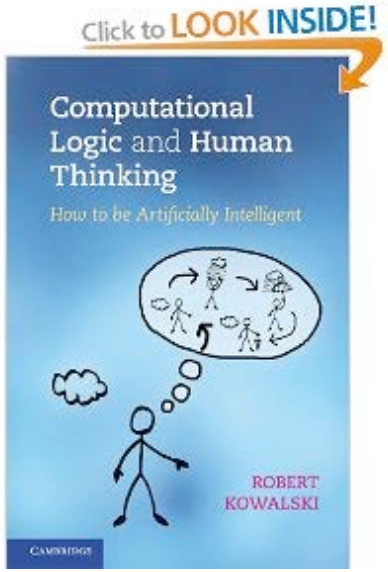
# A Connectionist implementation of ALP

# It can be difficult or impossible to put thoughts into words



# Conclusions

- The ALP agent model can unify
  - Logic
  - Production Systems
  - Probability
  - Decision Theory
  - Connectionism
- The ALP agent model can be used to improve human decision making.
- ALP clausal logic can be used to improve human communcation.


Click to LOOK INSIDE!
Computational Logic and Human Thinking
How to be Artificially Intelligent
ROBERT KOWALSKI
CAMBRIDGE

# The connection graph proof procedure

pre-computes links between the conditions and conclusions of clauses, together with their unifying substitutions.

Links can be activated, by performing resolution,
when the need arises.

Any strategy can be used for activating links,
including forwards and backwards reasoning.

Links that are activated frequently can be compiled into heuristic shortcuts, which achieve the same effects more directly.

## Connection graphs can combine logic, search, connectionism, learning and decision making

- Links can be weighted by statistics about how often they have contributed to successful outcomes in the past (and how likely they are to contribute in the future).
- Input observations and goals can be assigned different strengths (or utilities).
- The strength of observations and goals can be propagated through the graph in proportion to the weights on the links.
- Activating links with the currently highest weighted strengths implements a form of best-first search for a solution with highest expected utility, and is similar to the activation networks of Patie Maes.

## As Sherlock Holmes explained to Dr. Watson,

"In solving a problem of this sort, the grand thing is to be able to reason backward. That is a very useful accomplishment, and a very easy one, but people do not practise it much. In the everyday affairs of life it is more useful to reason forward, and so the other comes to be neglected. There are fifty who can reason synthetically for one who can reason analytically."

…….

"Most people, if you describe a train of events to them, will tell you what the result would be. They can put those events together in their minds, and argue from them that something will come to pass. There are few people, however, who, if you told them a result, would be able to evolve from their own inner consciousness what the steps were which led up to that result. This power is what I mean when I talk of reasoning backward, or analytically."

A theoretical framework for goal-based choice and for prescriptive analysis. Kurt A. Carlson & Chris Janiszewski & Ralph L. Keeney & David H. Krantz & Howard C. Kunreuther & Mary Frances Luce & J. Edward Russo & Stijn M. J. van Osselaer & Detlof von Winterfeldt. Market Lett (2008) 19:241–254.

"We view consumer preferences and consumer decisions as the output of goal pursuit. This departure from rational economic models allows us to characterize consumer behavior more fully.

For example, instead of assuming that consumer choices are the simple output of application of one's utility function to a set of known alternatives, with known consequences, we assume that choices result from consumers' attempt to satisfy numerous goals of different types (e.g., process, outcome, social), under a variety of constraints."

Suppose you receive an offer of a new job
when you are not looking for one.

It is tempting to limit your alternatives simply to
deciding between accepting or rejecting the offer.

But if you think about your broader goals,
then you might generate other alternatives,
like perhaps using the offer to negotiate
an improvement in your current employment.

(2) A new-born infant who, after commencement, is found abandoned in the United Kingdom shall, unless the contrary is shown, be deemed for the purposes of subsection (1) –

(a) to have been born in the United Kingdom after commencement; and

(b) to have been born to a parent who at the time of the birth was a British citizen or settled in the United Kingdom.

The logic of subsection 1.-(2) combines object language and meta-language:

*It shall be assumed that a person satisfies the conditions of subsection (1)*
*if the person is a new-born infant found abandoned in the United Kingdom after commencement*
*and it is not shown that the person does not satisfy the conditions of subsection (1)*