Herbert A. Simon's Legacy

Heuristics for Discovery in Cognitive Science

Pat Langley

Institute for the Study of Learning and Expertise

and

Center for the Study of Language and Information

Stanford University, Stanford, California

http://www.isle.org/~langley

langley@csli.stanford.edu

Heuristics and Scientific Discovery

Herbert Simon was fascinated by many phenomena, but two that drew his attention repeatedly were:

- the heuristic nature of human problem solving
- the processes of scientific reasoning and discovery

Thus, it seems appropriate to review Simon's career in terms of his personal heuristics for scientific research.

Moreover, it makes sense to illustrate these rules of thumb with examples from his own work on the discovery process.

Be Audacious

Tackle challenging problems that others have been reluctant to face or even admit are solvable.

- *Understand the cognitive and computational mechanisms that support the processes of scientific discovery.*

In 1966, Herb Simon published "Scientific Discovery and the Psychology of Problem Solving".

This radical paper set the agenda for research on computational scientific discovery for the next 35 years.

Ignore Discipline Boundaries

Become familiar with every field relevant to your research problem and incorporate the best ideas from each one.

- *To understand scientific discovery, borrow concepts not only from cognitive psychology and AI, but also from the history and philosophy of science.*

Herb Simon applied his Renaissance scholarship to his discovery research, as he did to many other scientific problems.

Moreover, he made his results accessible to members of all these communities by publishing in many literatures.

Use a Secret Weapon

Take advantage of metaphors and tools that you have mastered but that are not yet widely available.

- *Cast the discovery task in terms of heuristic search through a problem space controlled by a production system.*

Herb Simon repeatedly invoked the notion of heuristic search to model the discovery process, as to many other phenomena.

However, he was also ready to share his secret weapons with any who were willing to learn them.

Balance Theory and Data

Realize that scientific models must explain observations but also remain connected to existing knowledge.

- *Examine discoveries from the history of science that require computational explanation.*
- *Constrain these historical models using established knowledge about human cognition.*

Herb Simon's work on scientific discovery maintained a balance between theory and data, as did his other research efforts.

Satisfice

Address challenging problems but idealize them enough to make them tractable.

- *Focus on the discovery of descriptive laws from numeric data, producing BACON and its successors.*
- *Focus on discovery of simple structural models from qualitative data, producing STAHL and DALTON.*
- *Ignore issues of problem formulation, variable selection, and other aspects of scientific reasoning.*

However, Herb Simon always acknowledged the limits of a given idealization and the need for additional research.

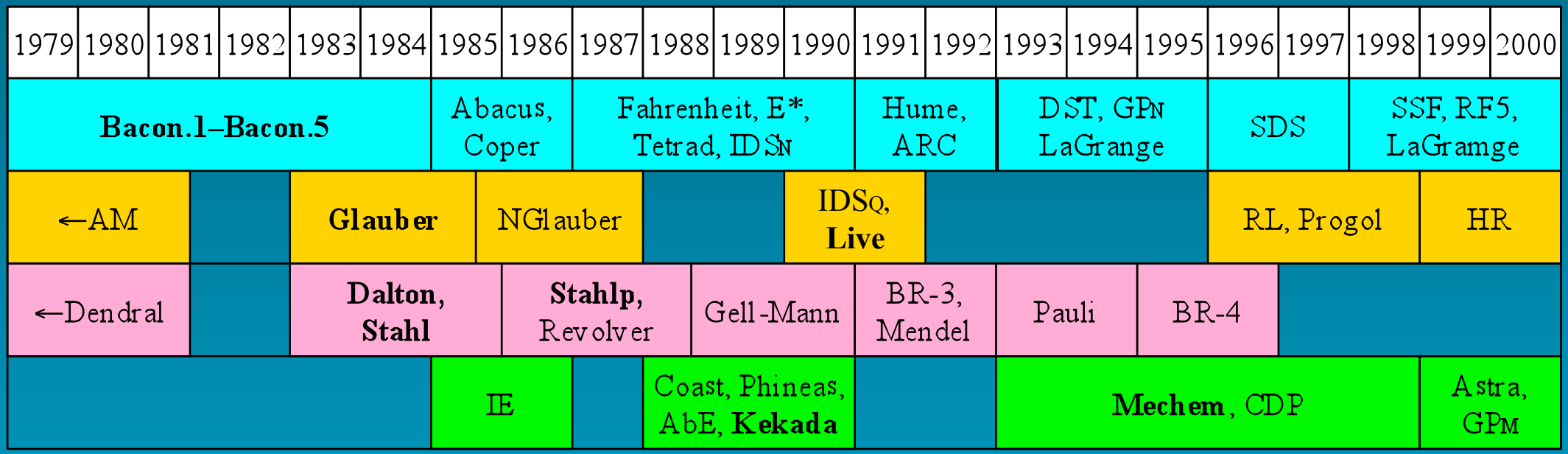
Persevere

Science is a gradual process. Build incrementally on your previous results, extending them to cover ever more phenomena.

- *Herb Simon and his colleagues worked steadily, for over two decades, to model the process of scientific discovery.*
- *Moreover, his research with Deepak Kulkarni on KEKADA itself modeled this central aspect of science.*

The resulting body of research helped change the face of cognitive science and clarified the computational nature of discovery.

Evolution of Research on Computational Scientific Discovery



Legend

Numeric laws	Qualitative laws	Structural models	Process models

Applications of Computational Discovery

Over the past decade, systems of this type have helped discover new knowledge in many scientific fields:

- stellar taxonomies from infrared spectra (Cheeseman et al., 1989)
- qualitative chemical factors in mutagenesis (King et al., 1996)
- quantitative laws of metallic behavior (Sleeman et al., 1997)
- qualitative conjectures in number theory (Colton et al., 2000)
- temporal laws of ecological behavior (Todorovski et al., 2000)
- reaction pathways in catalytic chemistry (Valdes-Perez, 1994, 1997)

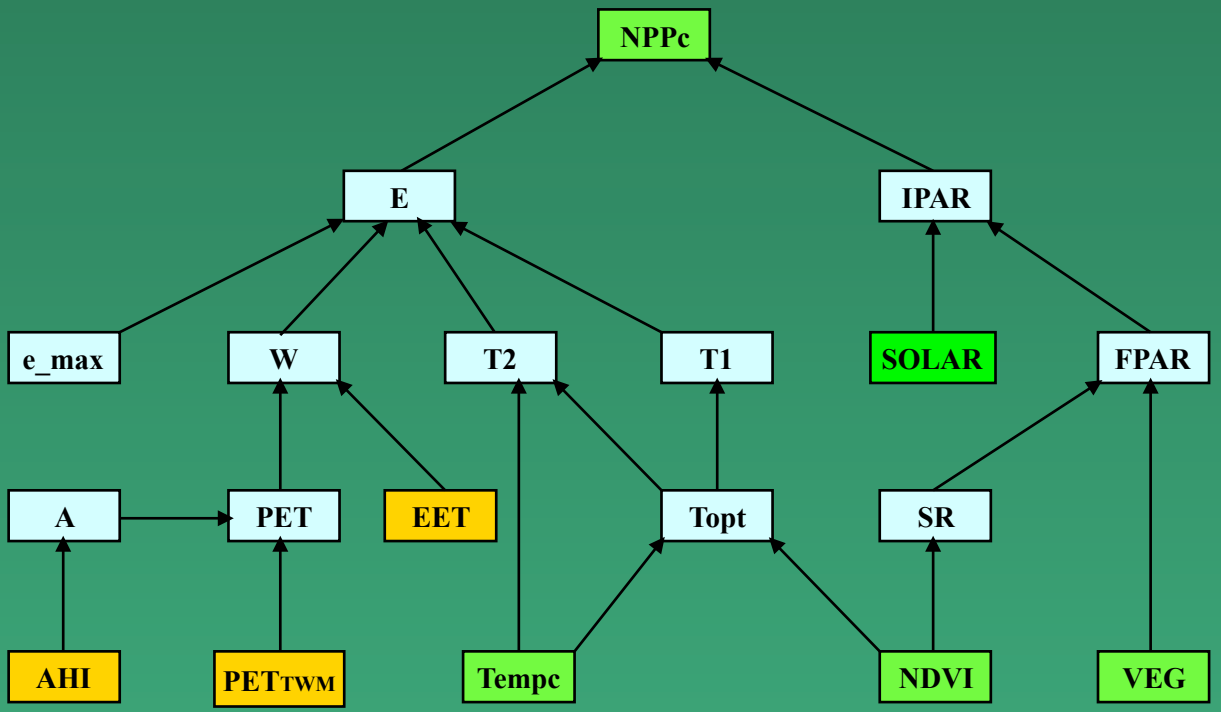
Each of these has led to publications in the refereed literature of the relevant scientific field.

Revising an Ecosystem Model

Given: A model of Earth's ecosystem (CASA) stated as difference equations that involve observable and hidden variables.

Given: Values of observable variables (rainfall, sunlight, NPP) as they change over both time and space.

Find: A revised ecosystem model with altered equations and/or parametric values that better fits the data.

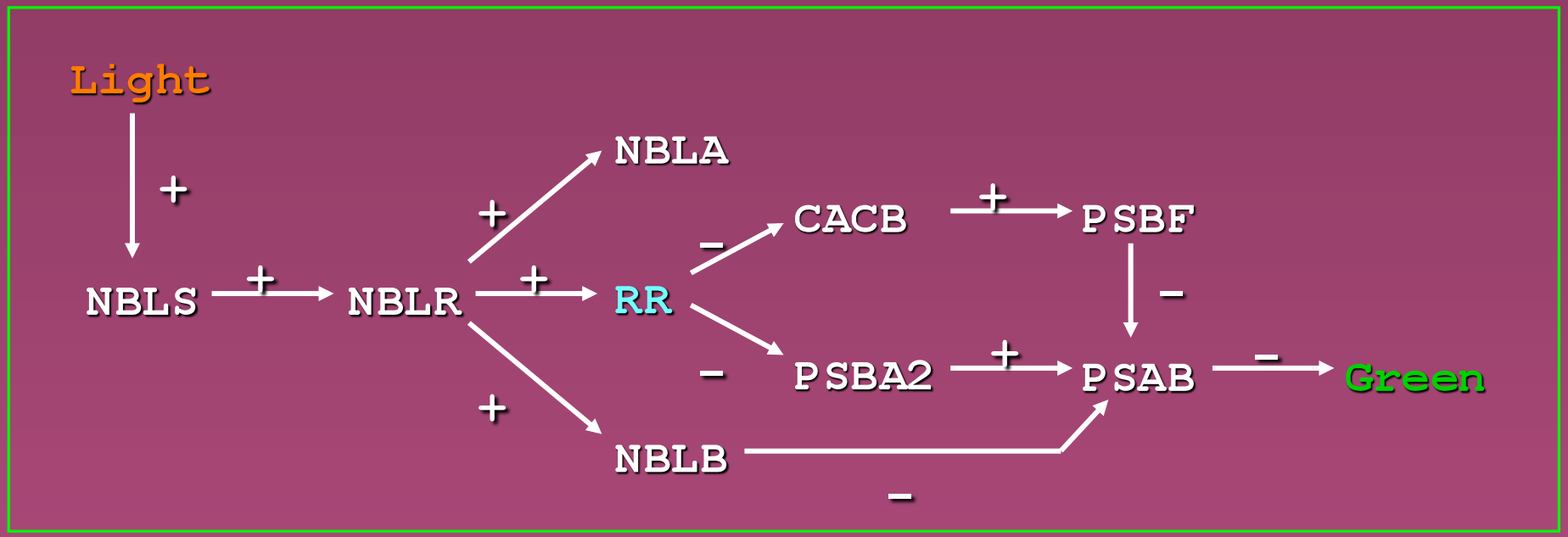


Revising Process Models of Photosynthesis

Given: Qualitative knowledge about reactions and regulations for Cyanobacteria in a high ultraviolet situation.

Given: Observed expression levels, over time, of the organism's genes under conditions of high light.

Find: A revised model with altered reactions and regulations that explains the expression levels and the bleaching process.



A Long-Term Goal

The ultimate challenge in discovery research is to model the behavior of a scientist who:

- Formulates the notion of satisficing in human decision making
- Co-invents list processing and heuristic search on computers
- Co-develops theories of human memory and problem solving
- Uses his theories to model discovery and other key phenomena
- Fosters a new field that acknowledges no discipline boundaries

We know some of this scientist's heuristics, and we have detailed records of his accomplishments, but the task remains daunting.

A Closing Quotation

We would like to imagine that the great discoverers, the scientists whose behavior we are trying to understand, would be pleased with this interpretation of their activity as normal (albeit high-quality) human thinking. . .

But science is concerned with the way the world is, not with how we would like it to be. So we must continue to try new experiments, to be guided by new evidence, in a heuristic search that is never finished but always fascinating.

Herbert A. Simon, Envoi to *Scientific Discovery*, 1987.