Guide (v1)

CONTENTS

The resources i have found so far!

1   Classical AI

2   Classical AI - Expert Systems (??)

Expert Systems: Principles and Programming, Fourth Edition

3   Qualitative Physics (knowledge representation)

TODO: find resources useful maybe for determining how people come up with 'naive' models to predict things, like fluids, etc. So we can determine what sorts of representations things in metaphors/analogies section have to build. I found this mentioned in the bibliography of /Artificial Intelligence: A modern approach.'

4   Metaphors/Analogies

4.1   **DONE** *Fluid Concepts and Creative Analogies*

4.2   **DONE** *Emotion Machine*

4.3   **STARTED** *LOGI*

4.4   *Godel Escher Bach: An Eternal Golden Braid ++*

4.5   *Metamagical Themas ++*

4.6   *Consciousness Explained ++*

4.7   *How The Mind Works ++*

4.8   *Women, Fire, and Dangerous Things*

4.9   *Perceptual Symbol Systems*

by Lawrence Barsalou's

4.10   *What is Thought?*

4.11   *The Symbolic Species - Terrence Deacon*

redundant?

5 Probability/Bayesian

5.1 *STARTED Data Analysis: A Bayesian Tutorial*

5.2 *Probability Theory: The Logic of Science*

5.3 *PLN (probabilistic logic network) book from OpenCog*

6 Memory/Case based reasoning

TODO: find case based reasoning resources. Visuo-spatial Working Memory MIT Encyclopedia of Cognitive Sciences

Sparse Distributed Memory shruti+variable binding+temporal synchronity holographic associative memory/geometric holo reduced rep/holo reduced representation

7 Heuristics and Causality and Utility

˜Eurisko,˜Scripts, Plans, Goals and Understandin

Causality: Models, Reasoning and Inference Choices, Values, and Frames

'Sources of Power' Simple Heuristics That Make Us Smart Bounded Rationality: The Adaptive Toolbox (intro to utility?) Judgment under uncertainty : heuristics and biases Heuristics and Biases: The Psychology of Intuitive Judgment

8 Evolutionary Psych

The Origins of Virtue

9 ++++Random Links

These are random links that i havent processed. General course

http://sites.google.com/site/narswang/home/agi-introduction

http://www.acceleratingfuture.com/michael/blog/category/ai/

Minsky is an actor in an artificial intelligence koan (attributed to his student, Danny Hillis) from the Jargon file:

In the days when Sussman was a novice, Minsky once came to him as he sat hacking at the PDP-6. "What are you doing?" asked Minsky. "I am training a randomly wired neural net to play Tic-tac-toe," Sussman replied. "Why is the net wired randomly?", asked Minsky. "I do not want it to have any preconceptions of how to play," Sussman said. Minsky then shut his eyes. "Why do you close your eyes?" Sussman asked his teacher. "So that the room will be empty." At that moment, Sussman was enlightened.

What I actually said was, "If you wire it randomly, it will still have preconceptions of how to play. But you just won't know what those preconceptions are." –Marvin Minsky

http://www.agiri.org/email/

http://www.sl4.org/archive//0512/13081.html Criticising is much, much easier than inventing; it takes far less time and knowledge to find a flaw in an existing proposal than invent a new one. When I was first learning AI I was constantly finding new things that seemed to work, but then as I improved and continued to research AGI these instances where rapidly outnumbered by finding out why things didn't work. Currently I manage to find something fatally wrong with >90% of the design ideas that I try within a few minutes to a few days, despite having what I'd consider fairly good intuition about what to investigate, without even

having an external review. Frankly the days when I could just make up something plausible and interesting sounding and proceed to implementation, as the vast majority of researchers do, were much more fun. But as I've often said before, it may be fun but it won't get you anywhere. Accurate criticism is vastly more useful than half-baked blue sky theorising. Most people do not have the time and the expertise to invent useful AGI/FAI theory, and in these cases finding concrete failings in existing work is a better allocation of effort than armchair speculation.

- Michael Wilson

Kanerva's 'sparse distributed memory' theory

http://www.sl4.org/archive/0512/13045.html It's true that AGI is somewhat all-or-nothing, but I don't think a simple estimate of completion time is much use at all. Arguably it's worse than useless as people often fixate on it and then decry you if you miss the deadline. I think to be useful you have to summarise you project plan into a set of major components, the key challenges for each, the dependencies between them, the resources assigned and a description of how the various capabilities your system should have will become available as you put the components together. Then you can label all that with confidence-bounded completion time estimates. Some people will probably still read it and reduce it down to 'they say they can do it in X years', but at least if you miss the deadline you can reference you project plan and show where you got things right and wrong, and meanwhile the people with a clue will be impressed that you made a serious effort to plan your project and justify your predictions. Personally I don't even have enough information to do this usefully yet, but I think I'm getting steadily closer to being able to.