Assessing the Intelligence of Cognitive Decathletes

Paul Cohen
USC Information Sciences Institute
Outline

• What is a cognitive decathlete?
  – Criteria for cognitive decathletes

• Discouraged, not dissuaded
  – How good tests can make the future bright
  – Attributes of good tests

• Designing the decathlon
  – What should the test test
  – Classes of problems ("events" in the decathlon)

• A specific proposal
The village all declared how much he knew; 'Twas certain he could write, and cipher too; Lands he could measure, terms and tides presage, And even the story ran that he could gauge. In arguing too, the parson owned his skill, For e'en though vanquished, he could argue still; While words of learned length and thundering sound Amazed the gazing rustics ranged around, And still they gazed, and still the wonder grew That one small head could carry all he knew.  
- Oliver Goldsmith
Two of my favorite cognitive decathletes

<table>
<thead>
<tr>
<th>Checklist:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Common sense knowledge</td>
<td>□</td>
</tr>
<tr>
<td>Facility with language</td>
<td>□</td>
</tr>
<tr>
<td>Learning new facts</td>
<td>□</td>
</tr>
<tr>
<td>Inference</td>
<td>□</td>
</tr>
<tr>
<td>Problem solving &amp; planning</td>
<td>□</td>
</tr>
<tr>
<td>Good manners, social skill</td>
<td>□</td>
</tr>
<tr>
<td>Sense of humor</td>
<td>□</td>
</tr>
<tr>
<td>Metacognitive knowledge</td>
<td>□</td>
</tr>
<tr>
<td>Preferences/purposive action</td>
<td>□</td>
</tr>
<tr>
<td>Creativity</td>
<td>□</td>
</tr>
<tr>
<td>Listen/perceive &amp; understand</td>
<td>□</td>
</tr>
<tr>
<td>Fill in missing bits</td>
<td>□</td>
</tr>
<tr>
<td>Ontology / classification</td>
<td>□</td>
</tr>
<tr>
<td>Memory and attention</td>
<td>□</td>
</tr>
</tbody>
</table>
Another favorite cognitive decathlete: Betty the New Caledonian Crow (courtesy Jackie Chappell, Alex Kacelnik)

**Intelligence . . . is the faculty of making artificial objects, especially tools to make tools. - Henri Bergson**

<table>
<thead>
<tr>
<th>Checklist:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common sense knowledge</td>
</tr>
<tr>
<td>Facility with language</td>
</tr>
<tr>
<td>Learning new facts</td>
</tr>
<tr>
<td>Inference</td>
</tr>
<tr>
<td>Problem solving &amp; planning</td>
</tr>
<tr>
<td>Good manners, social skill</td>
</tr>
<tr>
<td>Sense of humor</td>
</tr>
<tr>
<td>Metacognitive knowledge</td>
</tr>
<tr>
<td>Preferences/purposive action</td>
</tr>
<tr>
<td>Creativity</td>
</tr>
<tr>
<td>Listen/perceive &amp; understand</td>
</tr>
<tr>
<td>Fill in missing bits</td>
</tr>
<tr>
<td>Ontology / classification</td>
</tr>
<tr>
<td>Memory and attention</td>
</tr>
</tbody>
</table>
What *isn't* a cognitive decathlete?
Expert systems

*Ninety-five percent of all the decisions you'll ever make in your career would be made as well by any reasonably intelligent high school sophomore. But they'll pay you for the other five percent.*

- Marion Folsom
What isn't a cognitive decathlete?
Big Switch Systems
What isn't a cognitive decathlete?
Fakery

Where is your nose?
Entrant:...Oh, my nose? Why are you so interested?
Is it near your eyes?
Entrant:...Such is Human Perversity
Is your nose between your mouth and your eyes?
Entrant:...I can't tell you that
How many hands do I have?
Entrant:...I don't understand why you have, and furthermore, I don't really care
My daughter's name is Allegra. What is my daughter's name?
Entrant:...Is she a big part of your problem?
Architectures that might implement cognitive decathletes

"Modules"

Shared workspace

"Agents"

Tasks

Reasoning
Planning
Learning
Perception
Skill
…
Examples

Shakey the robot
Planning
action
perception
learning

Hearsay-II
Phonetics
Morphemics
Syntax
Semantics
Pragmatics

TRAINDS and TRIIPS
Language
Planning

Robocup robots
Perception
Planning
Teamwork
Learning
Information dependence
"The principle of continuously available output"

"Modules"

"Agents"

Components do their jobs better by sharing partial results – continuously available output (Norman & Bobrow) – re-entrant loops (Edelman) …
Criteria for cognitive decathletes

• Solve many classes of problems with one system of mutually-dependent components
• Solve many problems in each class
• Learn to solve new problems in each class
• When it makes sense to do so, use previously-learned information to solve new problems and to accelerate learning to solve new problems (but not all transfer is good transfer!)
Outline

• What is a cognitive decathlete?
  – Criteria for cognitive decathletes
• Discouraged, not dissuaded
  – How good tests can make the future bright
  – Attributes of good tests
• Designing the decathlon
  – What should the test test
  – Classes of problems ("events" in the decathlon)
  – Attributes of tests
• A specific proposal
Discouraged, not dissuaded

AI is still in its infancy. By that I mean that the incredibly powerful and ingenious AI theories, discoveries and creations of the past 50 years do not seem to me to have attained anything like the intelligence and capabilities of a six-year-old child. - Oliver Selfridge

- Discouraged: Specialization, the least publishable unit, tenure, low-hanging fruit, lack of imagination, funding agencies, and the difficulty of building large, integrated systems conspire against cognitive decathletes.

- Not dissuaded: Good problems and instruments produce good science, we have been asked to design a cognitive decathlon, DARPA can make the revolution happen, the future starts here!

I hereby offer to bet anyone a lobster dinner that by 2015 we will have a computer program capable of automatically reading at least 80 percent of the factual content across the entire English speaking web and placing those facts in a structured knowledge base. - Tom Mitchell
How to make the future bright 
(how tests can encourage the "right kind of AI")

• Computers are plastic, graduate students ingenious, nothing is proved by a program that solves exactly one kind of problem. Require one program to solve multiple kinds of problems using dependent, synergistic modules

• Many problems are trivially solved by "policies" which map from situations to best responses, or by more-or-less blind search. Emphasize interpretation and semantics over syntactic symbol manipulation, search, and instance-based policies

• Interpretation and semantics usually require tons of knowledge, so our programs won't do well. The solution is not to avoid problems that involve interpretation and semantics, but to accept knowledge-limited performance (even if poor) and emphasize learning (to increase knowledge)
How to make the future bright
(how tests can encourage the "right kind of AI")

• Discourage "divide and conquer"

An appropriate strategy...is divide and conquer
- study the dimensions of intelligence more or
less separately. We must be satisfied for a long
time to come with 'partial intelligence' in our
artifacts as a natural consequence of this
inevitable strategy. - Edward A. Feigenbaum
"Divide and conquer" or "Development"

A developmental strategy is to build more-or-less complete agents, each with several cognitive functions.

At first they aren't very capable.

By setting new problems we gradually make them more capable.
The developmental strategy

Build more-or-less complete agents, integrating several cognitive functions

Make the problems harder to make the agents more capable

Perhaps "divide and conquer" is not an inevitable strategy and we can do better than build partial intelligences.
Divide and conquer
Bad for scientists, bad for AIs

*It is not, truly speaking, the labour that is divided; but the men: divided into mere segments of men - broken into small fragments and crumbs of life, so that all the little piece of intelligence that is left in a man is not enough to make a pin, or a nail, but exhausts itself in making the point of a pin or the head of a nail.* - John Ruskin

"Agents"

Components do their jobs better by sharing information
Attributes of good tests

Organizational attributes

50-year technical and scientific goals

Frequent (e.g., annual) tests

Organizations to chart the way toward the goals via the tests
Attributes of good tests

50-year technical and scientific goals

- Frequent (e.g., annual) tests
- Organizations to chart the way toward the goals via the tests

Test important cognitive functions, particularly comprehension

- Automated scoring / continuously available test suites
- Simple success criteria
- Diagnosticity
- Specificity
- Transparency

Ample rope

Attributes of the tests themselves
Attributes of good tests

- 50-year technical and scientific goals
- Frequent (e.g., annual) tests
  - Organizations to chart the way toward the goals via the tests
- Test important cognitive functions, particularly comprehension
- Automated scoring / continuously available test suites
- Simple success criteria
- Diagnosticity
- Specificity
- Transparency
- Graduated series of challenges each just slightly out of reach
- Ample rope
- Monotonic, no "throwaways"
- Low cost of admission
- Come-as-you-are
- A popular problem / competition
- "Hearts and minds" attributes
Attributes of good tests

50-year technical and scientific goals

Frequent (e.g., annual) tests
Organizations to chart the way toward the goals via the tests

Test important cognitive functions, particularly comprehension
Automated scoring / continuously available test suites

Simple success criteria
Diagnostics
Specificity
Transparency

Graduated series of challenges each just slightly out of reach
Monotonic, no "throwaways"
A popular problem / competition

Development, not divide-and-conquer

Integrate AI technologies into more-or-less complete agents

In any given challenge, accept poor performance but insist on universal coverage

Paul Cohen. Do not distribute without permission
What about intelligence testing?

Machines already excel at Raven's-like and block-arrangement-like tasks ... knowledge free tasks that require speedy search through a space of configurations

But none I know of can interpret pictures from children's books.

*If we adopt IQ tests, adopt the parts that deal with interpretation, analogy, judgment, preferences, meaning*

A formal manipulator in mathematics often experiences the discomforting feeling that his pencil surpasses him in intelligence.

– Howard W. Eves
Outline

• What is a cognitive decathlete?
  – Criteria for cognitive decathletes
• Discouraged, not dissuaded
  – How good tests can make the future bright
  – Attributes of good tests
• Designing the decathlon
  – What should the test test?
  – What AI people think
  – What the rest of the world thinks
• A specific proposal
Which classes of problems?
What the Rest of the World Thinks

Preference, taste, motives in planning and action

*Good Living is an act of intelligence, by which we choose things which have an agreeable taste rather than those which do not.* -Brillat Savarin

*Does the proposed software have an adequate motivational structure? At the top level, we can use highly technical terms and ask if it gives a shit. ... It is not enough that the parent cares, for the child must care too. Our software should not be autistic offspring of us AI scientists and engineers.* -Oliver Selfridge
Which classes of problems?
What the Rest of the World Thinks

Judgment, ethics, wisdom

You're born with intelligence, but not with ethics. - Massad Ayoob

The fundamental cause of trouble in the world today is that the stupid are cocksure while the intelligent are full of doubt. - Bertrand Russell

I happen to feel that the degree of a person's intelligence is directly reflected by the number of conflicting attitudes she can bring to bear on the same topic. - Lisa Alther

Intelligence is not the same thing as wisdom. - Lao Tzu
Which classes of problems?
What the Rest of the World Thinks

Representing and acting in the physical world.
Physical skill, cognition, making representations

Skill is successfully walking a tightrope between the twin towers of New York's World Trade Center.
Intelligence is not trying. - Marilyn vos Savant

What is art? It is not just nature, it is nurtured nature. It is intelligence applied to what physical ability you have. - Rudolph Nuryev

There are painters who transform the sun into a yellow spot, but there are others who with the help of their art and their intelligence, transform a yellow spot into the sun. - Pablo Picasso

The essence of intelligence is skill in extracting meaning from everyday experience
Knowledge and intelligence

The knowledge principle (KP) A system exhibits intelligent understanding and action at a high level of competence primarily because of the specific knowledge that it can bring to bear: the concepts, facts, representations, methods, models, metaphors, and heuristics about its domain of endeavor.
- Lenat and Feigenbaum

I used to think she was quite intelligent ... because she knew quite a lot about the theater and plays and literature and all that stuff. If somebody knows quite a lot about those things, it takes you quite a while to find out whether they're really stupid or not. - J.D. Salinger
I am a great mayor; I am an upstanding Christian man; I am an intelligent man; I am a deeply educated man; I am a humble man. - Marion Barry
Which classes of problems?
What the Rest of the World Thinks

And for the average undergraduate…intelligence should be tested by a
These more commonsense notions of intelligence led to Gardner's Multiple Intelligences.
Seriously, though:
Which classes of problems?

<table>
<thead>
<tr>
<th>Perceive</th>
<th>Reason</th>
<th>Act</th>
<th>Learn</th>
</tr>
</thead>
</table>
| *Create internal representation, often by model selection* | *Forward*  
  *Predict*  
  *Classify*  
  *Backward*  
  *Explain*  
  *Plan*  
  *Maximize*  
  *Constraints*  
  *...* | *Control*  
  *Monitor*  
  *Kinematics*  
  *...* | *Model parameters*  
  *Models*  
  *Chunks*  
  *...* |
Outline

• What is a cognitive decathlete?
  – Criteria for cognitive decathletes

• Discouraged, not dissuaded
  – How good tests can make the future bright
  – Attributes of good tests

• Designing the decathlon
  – What should the test test?
  – Classes of problems ("events" in the decathlon)

• A specific proposal
How did I get so smart?

Computers can't do any of these things well, never mind all of them in concert. However will they *make contact* with our knowledge?
How can my program get smart?

Fake vision and other perceptual modes

Real speech and reading

Fake physical actions, real experiments

AI

Fake vision and other perceptual modes

It is easier for machines to make contact with artificial environments than with natural ones

"Alice, try turning left"

"What is Susie having to eat?"
Events in a cognitive decathlon

Perception

What is happening on a physical level; give an intentional gloss; give an ethical gloss

Reasoning

Why is it happening, what will happen next? Do the actors like each other?

Planning and Action

Where should go to avoid Is there anywhere to hide?

Learning (see next slides)

Heider and Simmel, 1944
Learning

What is happening?
Does Red like Blue?
Does Blue help Red?
Is Blue hiding from Red?
What would you do to Red?

Is Red the same person as in the previous movie?
Has Red's relationship to Blue changed?
Can blobs change their behaviors?
A cognitive decathlete can:

• Solve many classes of problems with one system of mutually-dependent components – *at least: perception of facts, inference of intentions, prediction, explanation, judgments of utility, planning, social relations, ethical relations*

• Solve many problems in each class – *as many as you want!*

• Learn to solve new problems in each class – *which patterns of movement indicate intentional states? Which intentional states indicate future actions? Which actions indicate interpersonal relations?*

• When it makes sense to do so, use previously-learned information to solve new problems and to accelerate learning to solve new problems (but not all transfer is good transfer!)
Metrics: Gold Standards

• Compare answers to all levels of all questions (factual, intentional, ethical) with those of children; correlate with developmental stages, e.g., Kohlberg questions.
• Compare AI's planning to children's planning
• Compare AI's explanations to children's explanations
• Compare AI's predictions to children's predictions
• Compare children's learning to AI's learning

• And so on: Children are the gold standard, and the gold standard is easy to achieve within this simple experimental paradigm.
Who cares about children as a gold standard?

• I do, because it gauges the "level" of AI intelligence, but exactly the same paradigm works for movies like this, with military officers as the gold standard.

What is happening? Why is it happening? What might happen next? What can you do to avoid bad outcomes? What are the spatial / intentional / causal relationships between actors....
Cognitive Decathlon Events

1. Interpret movie/game actions
   - Physical interpretation
   - Intentional interpretation
   - Utility/ethical/taste interpretation
2. Predict what will happen
3. Explain what happened and one's predictions; build causal models of objects and agents in the world
4. Infer relations – spatial, intentional, interpersonal…
   - Can A see B; does A help B; does A like B…
5. Plan – what would you do to achieve/avoid …
6. Act – Try it and see what happens
7. Learn to improve performance on all of the above
8. Communicate about all of the above (in formal or natural language) which makes learning from instruction possible
Conclusion

A cognitive decathlete can:

• Solve many classes of problems with one system of mutually-dependent components
• Solve many problems in each class
• Learn to solve new problems in each class
• When it makes sense to do so, use previously-learned information to solve new problems and to accelerate learning

• A good test
  • Is one of a graduated series of challenges, passable today (albeit with effort), on a path to human-like intelligence, providing ample rope, diagnosticity, reliability, defense against divide and conquer
  • Stresses interpretation, understanding, semantics
  • Requires learning and cognitive development

• I think we gotta good one!