I agree w/ Joscha's caution about discrimination tasks: They can be
often be solved rather well, but in devious ways, by statistical
supervised learning algorithms.

The attached picture will get our current best statistical methods to tell us that it sees four to six people.
A really good system might recognize that two move forward and two in the opposite direction, and that the
latter ones have full bags.
I don't think that there is a system that would tell us that there is probably a store in the direction the bagless
people are going.
Most of all, current systems probably won't figure out that there must be a sniper behind that wall.
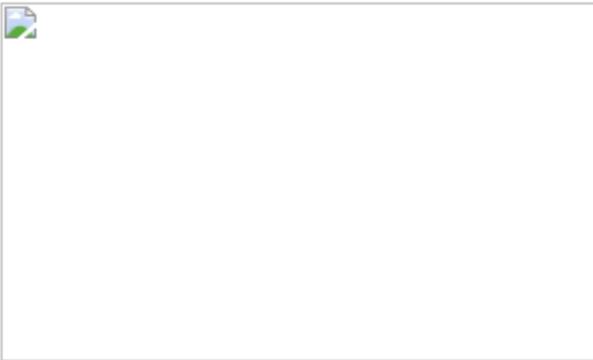
People don't stop at matching patterns; they construct a conceptual world view, and integrate what they see, hear
and read into it. IMHO, this is what an AI challenge needs to be about: use any kind of information, and integrate
it into a deeper, growing and dynamic understanding of the world.

One way of doing that might be to let it re-tell the story of a movie we show it. We will have to make sure that
the particular movie is unknown to the system (it is likely going to be trained on many annotated movies), which
could be achieved by picking one (or rather, several, with different degrees of difficulty) that has not aired yet
when the submission is made. We might also put limits on the memory footprint of the system, to make sure that
it does not memorize existing stuff too literally, but is forced to make inferences. The whole thing could be a
competition, where the performance is compared to children of different ages, using a mixed jury including
developmental psychologists, computer scientists and screenwriters.

Prices could be given to contributions that match the performance of a 3yr old, 6yr old, and adult, for a challenge
of five movies unknown to the submitters.

Suppose you pose a linguistic
discrimination task of some sort -- and a supervised learning
algorithm, trained on a mass of data, can solve it with 97% accuracy.

I don't think that Bob Berwick knew Deep Speech yet, a system that Andrew Ng built for Baidu earlier this year. It does not do any Fourier transform on auditory data, but only raw convolutional networks. It works on phone audio, noisy streets, in conference rooms etc., and it matches what it hears to phonemes, which are then mapped to language on a separate layer. According to the publication, it outperforms humans at this.

Bob told as that he knows how language itself works, i.e. not the mapping from sound to phonemes, but the much more interesting part behind that. What is missing, from an AI perspective, is the link between language and the simulated inner world in our minds. I think that AI is still quite a way from recreating our imagination (and Noam is skeptical that it will ever happen, if I understand him correctly). For a Chomsky challenge, perhaps we would need to find a task that requires deep language models without necessarily requiring deep understanding. I am sure that the linguists have much better ideas than me, of course.

– Joscha