Expression of Interest

Creating Ethical Robots: A Grand Challenge

James Gips
Egan Professor of Computer Science
Boston College

gips@bc.edu
http://www.cs.bc.edu/~gips

I am interested in formulating the problem of the creation of ethical robots as a Grand Challenge, especially as defined by the UK Computing Research Committee:

(from http://www.nesc.ac.uk/esi/events/Grand_Challenges/criteria.html)

Grand Challenges for Computing Research

Sponsored by the UK Computing Research Committee, with support from EPSRC and NeSC

CRITERIA OF MATURITY FOR A GRAND CHALLENGE (2002)

The chief purpose of the formulation and promulgation of a grand challenge is the advancement of science. A grand challenge represents a commitment by a significant scientific community to work together towards a common goal, agreed to be valuable and achievable within a predicted timescale. The challenge is formulated by the scientists themselves as a focus for the research that they wish to pursue in any case. It is independent of any political initiatives or prior allocation of special funding. It may involve a thousand man-years of research effort, drawn from many countries and spread over ten years or more. The main barrier to its faster progress is often the shortage of dedicated scientists of the right calibre and speciality. An opportunity for a

grand challenge arises only rarely in the history of science, when a branch of study first reaches an adequate level of maturity to predict and plan the direction of future progress.

The purpose of the list given below is to clarify the criteria of maturity as applied to a proposed scientific challenge. The suggested criteria concentrate on those aspects that contribute towards the primary goal of a grand challenge, which is the advancement of science. It is this that distinguishes a grand challenge from the many other worthy kinds of challenge, formulated to contribute to economic, political, military or other goals of society. No challenge, however grand or feasible or otherwise desirable, should be expected to meet all the criteria. The order of the criteria is not significant.

- It arises from scientific curiosity about the foundation, the nature or the limits of a scientific discipline.
- It gives scope for engineering ambition to build something that has never been seen before.
- It will be obvious how far and when the challenge has been met (or not).
- It has enthusiastic support from (almost) the entire research community, even those who do not participate and do not benefit from it.
- It has international scope: participation would increase the research profile of a nation.
- It is generally comprehensible, and captures the imagination of the general public, as well as the esteem of scientists in other disciplines.
- It was formulated long ago, and still stands.
- It promises to go beyond what is initially possible, and requires development of understanding, techniques and tools unknown at the start of the project.
- It calls for planned co-operation among identified research teams and communities.
- It encourages and benefits from competition among individuals and teams, with clear criteria on who is winning, or who has won.
- It decomposes into identified intermediate research goals, whose achievement brings scientific or economic benefit, even if the project as a whole fails.
- It will lead to radical paradigm shift, breaking free from the dead hand of legacy.
- It is not likely to be met simply from commercially motivated evolutionary advance.

THE GRANDNESS OF A GRAND CHALLENGE

The tradition of Grand Challenges is common in many branches of Science. If you want to know whether a challenge qualifies for the title 'Grand', compare it with

Put a man on the moon within ten years	(accomplished, 1960s)
Cure cancer within in ten years	(failed, 1970s)
Prove Fermat's last theorem	(accomplished)
Map the Human Genome	(accomplished)
Map the Human Proteome	(too difficult for now)
Find the Higgs boson	(under investigation)
Find Gravity waves	(under investigation)

Unify the four forces of Physics   (under investigation)

Complete Hilbert's programme for mathematical logic (almost complete)

In Computer Science, the following are listed not as recommendations but as examples that may be familiar from the past.

Prove that P is not equal to NP   (open)

The Turing test   (inactive)

The verifying compiler   (abandoned, 1970s)

A championship chess program   (completed)

A GO program at professional standard (too difficult)

Automatic translation from Russian to English (failed, 1960s)

A mathematical model of the evolution of the web (new)

A wearable computer serving as a guide dog for the blind (new)

These challenges are motivated primarily by scientific curiosity about the ultimate scope and limitations of computers, or by engineering ambition to construct something that has never been built before. This is a criterion which distinguishes a grand challenge from the many other challenges that have been proposed and accepted by computer scientists, ones that are motivated primarily by goals that have been set by society, often economic, political, or military goals. The adoption and promotion of a grand challenge is not intended to compete with this more familiar kind of challenge.