

# Mapping the Landscape of Human-Level Artificial General Intelligence

## Sam S. Adams

IBM Research  
ssadams@us.ibm.com

## Itamar Arel

University of Tennessee, Knoxville  
itamar@ece.utk.edu

## Joscha Bach

Humboldt University of Berlin  
joscha.bach@hu-berlin.de

## Robert Coop

University of Tennessee, Knoxville  
rcoop@utk.edu

## Rod Furlan

Singularity University  
rfurlan@gmail.com

## Ben Goertzel

Novamente LLC and Xiamen University  
ben@goertzel.org

## J. Storrs Hall

Institute for Molecular Manufacturing  
josh@foresight.org

## Alexei Samsonovich

George Mason University  
samsonovich@cox.net

## Matthias Scheutz

University of Indiana  
mscheutz@indiana.edu

## Matthew Schlesinger

Southern Illinois University, Carbondale  
matthews@siu.edu

## Stuart C. Shapiro

SUNY Buffalo  
shapiro@buffalo.edu

## John Sowa

Vivomind  
sowa@bestweb.net

## Abstract

We present the broad outlines of a roadmap toward human-level artificial general intelligence (henceforth, AGI). We begin by discussing AGI in general, adopting a pragmatic goal for its attainment and a necessary foundation of characteristics and requirements. An initial capability landscape will be presented, drawing on major themes from developmental psychology and illuminated by mathematical, physiological and information processing perspectives. The challenge of identifying appropriate tasks and environments for measuring AGI will be addressed, and seven scenarios will be presented as milestones suggesting a roadmap across the AGI landscape along with directions for future research and collaboration.

**Keywords:** AI, AGI, Artificial Intelligence, Artificial General Intelligence, Roadmap, Human-level

## Introduction

This paper is the result of an ongoing collaborative effort by the co-authors, preceding and during the “AGI Roadmap Workshop” held at the University of Tennessee, Knoxville in October 2009, and from many continuing discussions since then. Some of the ideas also trace back to discussions held during two workshops on “Evaluation and Metrics for Human Level AI” organized by John Laird and Pat Langley (one in Ann Arbor in late 2008, and one in

Tempe in early 2009). Some of the conclusions of the Ann Arbor workshop were recorded in (Laird et al. 2009). Inspiration was also obtained from discussion at the “Future of AGI” post-conference workshop of the AGI-09 conference, triggered by Itamar Arel’s presentation on the “AGI Roadmap” theme (Arel 2009); and from an earlier article on AGI road-mapping (Arel and Livingston 2009).

Of course, this is far from the first attempt to plot a course toward human-level AGI: arguably this was the goal of the founders of the field of Artificial Intelligence in the 1950s, and has been pursued by a steady stream of AI researchers since, even as the majority of the AI field has focused its attention on more narrow, specific subgoals. The ideas presented here build on the ideas of others in innumerable ways, but to review the history of AI and situate the current effort in the context of its predecessors would require a much longer article than this one. Thus we have chosen to focus on the results of our AGI roadmap discussions, acknowledging in a broad way the many debts owed to many prior researchers. References to the prior literature on evaluation of advanced AI systems are given in (Laird et al. 2009) and (Goertzel and Bugaj 2009), which may in a limited sense be considered prequels to this paper.

We begin by discussing AGI in general and adopt a pragmatic goal for measuring progress toward its attainment. We also adopt, as a provisional starting point, a

slightly modified version of the characteristics and requirements for AGI proposed in (Laird and Wray 2010), upon which we will later construct a number of specific scenarios for assessing progress in achieving AGI. An initial capability landscape for AGI will be presented, drawing on major themes from developmental psychology and illuminated by mathematical, physiological and information processing perspectives. The challenge of identifying appropriate tasks and environments for measuring AGI will be taken up. Several scenarios will be presented as milestones outlining a roadmap across the AGI landscape and directions for future work and collaboration will conclude the paper.

## The Goal: Human-Level General Intelligence

Simply stated, the goal of AGI research as considered here is the development and demonstration of systems that exhibit the broad range of general intelligence found in humans. This goal of developing AGI echoes that of the early years of the Artificial Intelligence movement, which after many valiant efforts largely settled for research into “narrow AI” systems that could demonstrate or surpass human performance in a specific task, but could not generalize this capability to other types of tasks or other domains.

A classic example of this narrow AI approach was IBM’s DeepBlue system (Campbell, Hoane, and Hsu 2002), which successfully defeated world chess champion Gary Kasparov but could not readily apply that skill to any other problem domain without substantial human re-programming. In early 2011, IBM’s WATSON question answering system (Ferrucci 2010) dramatically defeated two all time champions in the quiz show *Jeopardy!*, but having never “personally” visited Chicago’s O’Hare and Midway airports, fumbled on a question that any human frequent flier would have known. To apply the technology underlying WATSON to another domain, such as healthcare or call center support, would require not merely education of the AI system, but significant reprogramming and human scoring of relevant data -- the analogue of needing to perform brain surgery on a human each time they need to confront a new sort of task. As impressive as these and other AI systems are in their restricted roles, they all lack the basic cognitive capabilities and common sense of a typical five year old child, let alone a fully educated adult professional.

Given the immense scope of the task of creating AGI, we believe the best path to long term success is collaboration and coordination of the efforts of multiple research groups. A common goal and a shared understanding of the landscape ahead of us will be crucial to that success, and it was the aim of our workshop to make substantial progress in that direction.

## A pragmatic goal for AGI

The heterogeneity of general intelligence in humans makes it practically impossible to develop a comprehensive, fine-grained measurement system for AGI. While we encourage research in defining such high-fidelity metrics for specific capabilities, we feel that at this stage of AGI development a pragmatic, high level goal is the best we can agree upon.

Nils Nilsson, one of the early leaders of the AI field, stated such a goal in the 2005 AI Magazine article “Human-Level Artificial Intelligence? Be Serious!” (Nilsson, 2005):

“I claim achieving real human-level artificial intelligence would necessarily imply that most of the tasks that humans perform for pay could be automated. Rather than work toward this goal of automation by building special-purpose systems, I argue for the development of general-purpose, educable systems that can learn and be taught to perform any of the thousands of jobs that humans can perform. Joining others who have made similar proposals, I advocate beginning with a system that has minimal, although extensive, built-in capabilities. These would have to include the ability to improve through learning along with many other abilities.”

Many variant approaches have been proposed for achieving such a goal, and both the AI and AGI communities have been working for decades on the myriad subgoals that would have to be achieved and integrated to deliver a comprehensive AGI system. But aside from the many technological and theoretical challenges involved in this effort, we feel the greatest impediment to progress is the absence of a common framework for collaboration and comparison of results. The AGI community has been working on defining such a framework for several years now and, as mentioned earlier, this present effort on AGI road-mapping builds on the shoulders of much prior work.

## Characteristics and Requirements for AGI

As a starting point for our road-mapping effort, we have adopted a slightly modified version of Laird and Wray’s “Cognitive Architecture Requirements for Achieving AGI” (Laird and Wray 2010). Their outline of eight characteristics for environments, tasks and agents and twelve requirements for general cognitive architectures provide a level foundation for the comparison of appropriate scenarios for assessing AGI systems. We stress, however, that in our perspective these requirements should not be considered as final or set in stone, but simply as a convenient point of departure for discussion and collaboration. As our understanding of AGI improves through ongoing research, our understanding of the associated requirements is sure to evolve. In all probability, each of the many current AGI

research paradigms could be used to spawn a slightly different requirements list, but we must start somewhere if we are to make progress as a community.

To test the capability of any AGI system, the characteristics of the intelligent agent and its assigned tasks within the context of a given environment must be well specified. Failure to do this may result in a convincing demonstration, but make it exceedingly difficult for other researchers to duplicate experiments and compare and contrast alternative approaches and implementations.

- C1. The environment is complex, with diverse, interacting and richly structured objects.
- C2. The environment is dynamic and open.
- C3. Task-relevant regularities exist at multiple time scales.
- C4. Other agents impact performance.
- C5. Tasks can be complex, diverse and novel.
- C6. Interactions between agent, environment and tasks are complex and limited.
- C7. Computational resources of the agent are limited.
- C8. Agent existence is long-term and continual.

figure 1. Characteristics for AGI Environments, Tasks, and Agents

The characteristics shown in figure 1 provide the necessary (if not sufficient) degrees of dynamism and complexity that will weed out most “narrow AI” approaches at the outset, while challenging researchers to continually consider the larger goal of AGI during their work on subsystems and distinct capabilities. We have added the requirement for openness of the environment (C2), i.e., the agent should not be able to rely on a fixed library of objects, relations and events. We also require objects to have an internal structure that requires complex, flexible representations (C1).

There are nearly as many different AGI architectures as there are researchers in the field. If we are ever to be able to compare and contrast systems, let alone integrate them, a common set of architectural features must form the basis for that comparison.

Figure 2 reprises Laird and Wray’s requirements for general cognitive architectures and provides a framework for that basis. We have modified requirement R0, which in Laird and Wray reads “Fixed structure for all tasks”, to emphasize that the system may grow and develop over time, perhaps changing its structure dramatically – but these changes need to be effected by the agent itself, not by the intervention of the programmer. We anticipate further additions and changes to these lists over time as the community converges in experience with various cognitive

architectures and AGI systems, but we have adopted them so we can progress as a community to our larger, shared goals.

- R0. New tasks do not require re-programming of the agent
- R1. Realize a symbol system
- Represent and effectively use:
  - R2. Modality-specific knowledge
  - R3. Large bodies of diverse knowledge
  - R4. Knowledge with different levels of generality
  - R5. Diverse levels of knowledge
  - R6. Beliefs independent of current perception
  - R7. Rich, hierarchical control knowledge
  - R8. Meta-cognitive knowledge
- R9. Support a spectrum of bounded and unbounded deliberation
- R10. Support diverse, comprehensive learning
- R11 Support incremental, online learning

figure 2: Cognitive Architecture Requirements for AGI

## Current Challenges

In addition to providing a framework for collaborative research in AGI, Laird and Wray identified two significant challenges:

“...one of the best ways to refine and extend these sets of requirements and characteristics is to develop agents using cognitive architectures that test the sufficiency and necessity of all these and other possible characteristics and requirements on a variety of real-world tasks. One challenge is to find tasks and environments where all of these characteristics are active, and thus all of the requirements must be confronted. A second challenge is that the existence of an architecture that achieves a subset of these requirements, does not guarantee that such an architecture can be extended to achieve other requirements while maintaining satisfaction of the original set of requirements.” (Laird and Wray 2010)

Our 2009 AGI road-mapping workshop took up the first challenge of finding appropriate tasks and environments to assess AGI systems, while the second challenge will be more appropriately handled by individual research efforts. We also added a third challenge, that of defining the landscape of AGI, in service to the AGI road-mapping effort that has been underway for several years.

The balance of this paper will deal with these two challenges in reverse order. First, we will provide an initial landscape of AGI capabilities based on developmental psy-

chology and underpinned by mathematical, physiological and information processing perspectives. Second, we will discuss the issues that arise in selecting and defining appropriate tasks and environments for assessing AGI, and finally, we will present a number of scenarios and directions for future work and collaboration.

### **Challenge: Mapping the Landscape**

There was much discussion both in preparation and throughout our AGI Roadmap Workshop on the process of developing a roadmap. A traditional highway roadmap shows multiple driving routes across a landscape of cities and towns, natural features like rivers and mountains, and political features like state and national borders. A technology roadmap typically shows a single progression of developmental milestones from a known starting point to a desired result. Our first challenge in defining a roadmap for achieving AGI was that we initially had neither a well defined starting point nor a commonly agreed upon target result. The history of both AI and AGI is replete with this problem, which is somewhat understandable given the breadth and depth of the subjects of both human intelligence and computer technology. We made progress by borrowing more metaphors from the highway roadmap, deciding to first define the landscape for AGI and then populate that landscape with milestones that may be traversed via multiple alternative routes.

The final destination, full human-level artificial general intelligence, encompasses a system that could learn, replicate (and possibly exceed) human level performance in the full breadth of cognitive and intellectual abilities. The starting point, however, was more problematic, since there are many current approaches to achieving AGI that assume different initial states. We finally settled on a developmental approach to the roadmap, following human cognitive development from birth through adulthood.

The various scenarios for assessing AGI that had been submitted by participants prior to the workshop were then arrayed as milestones between these two endpoints, without any specific routes between them aside from their relative requirements for increasing levels of human cognitive development.

### **Top Down: Characterizing Human Cognitive Development**

The psychological approach to intelligence encompasses a broad variety of sub-approaches rather than presenting a unified perspective. Viewed historically, efforts to conceptualize, define, and measure intelligence in humans reflect a distinct trend from general to specific (Gregory 1996), much like the history of AI.

Early work in defining and measuring intelligence was heavily influenced by Spearman, who in 1904 proposed the psychological factor  $g$  (for general intelligence). Spearman argued that  $g$  was biologically determined, and represented the overall intellectual skill level of an individual. A related advance was made in 1905 by Binet and Simon, who developed a novel approach for measuring general intelligence in French schoolchildren. A unique feature of the Binet-Simon scale was that it provided comprehensive age norms, so that each child could be systematically compared with others across both age and intellectual skill level. In 1916, Terman introduced the notion of an intelligence quotient or  $IQ$ , which is computed by dividing the test-taker's mental age (i.e., their age-equivalent performance level) by their physical or chronological age.

In subsequent years, psychologists began to question the concept of intelligence as a single, undifferentiated capacity. There were two primary concerns. First, while performance within an individual across knowledge domains is somewhat correlated, it is not unusual for skill levels in one domain to be considerably higher or lower than in another (i.e., intra-individual variability). Second, two individuals with comparable overall performance levels might differ significantly across specific knowledge domains (i.e., inter-individual variability).

These issues helped to motivate a number of alternative theories, definitions, and measurement approaches, which share the idea that intelligence is multifaceted and variable both within and across individuals. Of these approaches, a particularly well-known example is Gardner's theory of multiple intelligences (Gardner 1999), which proposes eight distinct forms or types of intelligence: (1) linguistic, (2) logical-mathematical, (3) musical, (4) bodily-kinesthetic, (5) spatial, (6) interpersonal, (7) intrapersonal, and (8) naturalist. Gardner's theory suggests that each individual's intellectual skill is represented by an intelligence profile, a unique mosaic or combination of skill levels across the eight forms of intelligence.

While Gardner's theory has had significant impact within the field of adult intelligence, it has had comparatively less influence on the study of intellectual development in children. Instead, researchers in the field of cognitive development seek to (1) describe processes of intellectual change, while (2) identifying and explaining the underlying mechanisms (both biological and environmental) that make these changes possible. Contemporary theories of cognitive development are very diverse and defy simple systematization and a thorough treatment of the field would take us too far from our focus. However, two major schools of thought, those of Piaget and Vygotsky, will serve as axes for our AGI landscape.

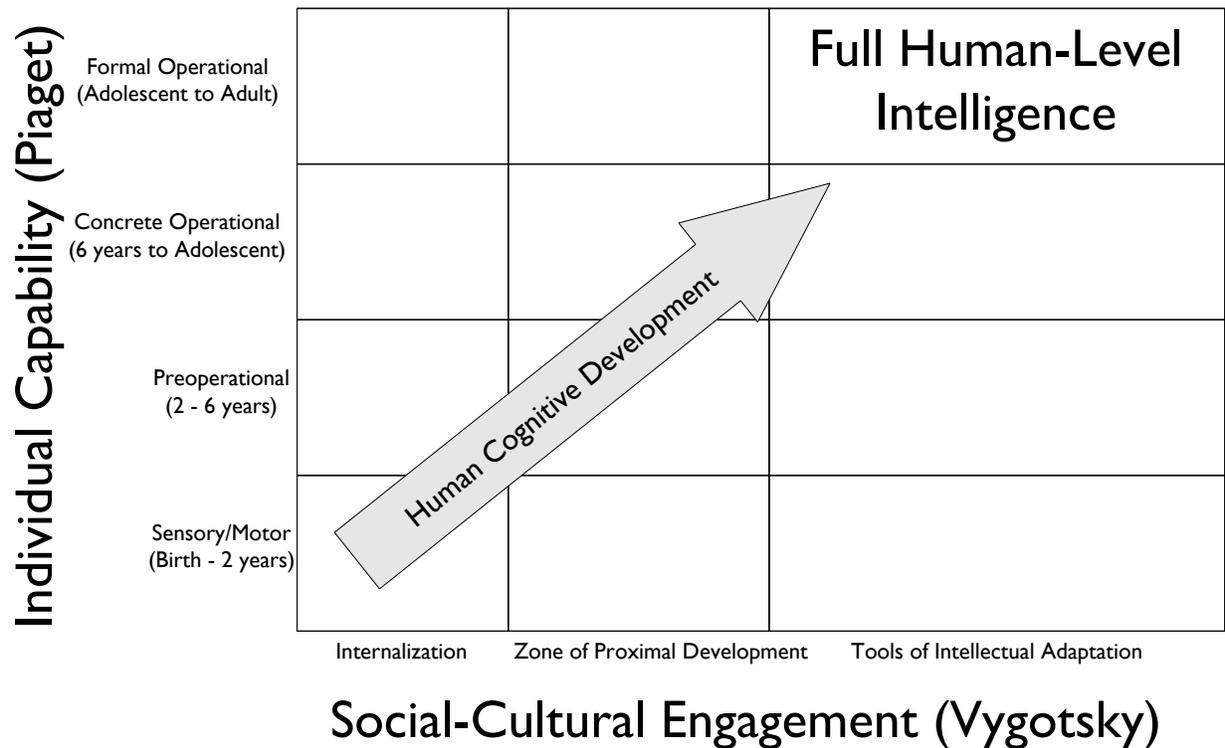


figure 3. The Landscape of Human Cognitive Development

### Piaget's Theory

In his classic work that founded the science of cognitive development, Piaget proposed that humans develop through four qualitatively-distinct stages of development (Piaget 1953).

First, during the sensorimotor stage (0-2 years), infants acquire a rich repertoire of perceptual and motor skills (e.g., reaching, grasping, crawling, walking, etc.). This stage includes a number of major milestones, such as the ability to search for hidden objects and the ability to use objects as simple tools.

Second, infants enter the preoperational stage (2-6 years) as they acquire the capacity to mentally represent their experiences (e.g., memory, mental images, drawing, language, etc.), but lack the ability to systematically coordinate these representations in a logically consistent manner.

During the next stage (6 years to adolescence), children at the concrete operational level master basic elements of logical and mathematical thought, including the ability to reason about classes and categories, as well as numerical operations and relations.

The final stage of development, formal operations, begins in adolescence and includes the use of deductive logic, combinatorial reasoning, and the ability to reason about hypothetical events.

### Vygotsky's Theory

In contrast to Piaget's view which focuses on the individual, Vygotsky's classic theory of cognitive development

emphasizes the socio-cultural perspective (Vygotsky 1986). Vygotsky's theory not only highlights the influence of the social environment, but it also proposes that each culture provides a unique developmental context for the child. Three fundamental concepts from this theory are (1) internalization, (2) zone of proximal development, and (3) tools of intellectual adaptation.

First, Vygotsky proposed that the capacity for thought begins by acquiring speech (i.e., thinking "out loud"), which gradually becomes covert or internalized.

Second, he emphasized that parents, teachers, and skilled peers facilitate development by helping the child function at a level just beyond what they are capable of doing alone; he referred to this cognitive space as the zone of proximal development.

Third, Vygotsky also stressed that each child "inherits" a unique set of objects, ideas, and traditions that guide learning (e.g., books, calculators, computers, etc.). These tools of intellectual adaptation not only influence the pattern of cognitive development, but also serve as constraints on the rate and extent of development.

### Surveying the Landscape of Human Intelligence

While many consider the views of Piaget and Vygotsky to be at odds because of their different focus, we consider them to be complementary, partial descriptions of the same development process. Piaget's approach focused on the stages of cognitive development of an individual child, while Vygotsky considered the same development within

the context of social interaction with other humans with access to and training in the use of cultural artifacts like tools, languages, books, and a shared environment. By placing the developmental stages of each theory on opposing axes (figure 3), we can outline the landscape of human cognitive development and provide a structure for the placement of milestones along the road to AGI.

### Bottom Up: Substrata of the AGI Landscape

Our ultimate goal as a community is to create functional computer systems that demonstrate human-level general intelligence. This lofty aim requires that we deeply understand the characteristics of human cognitive behavior and development outlined above, as well as implement that understanding in a non-biological substrate, the modern digital computer. To accomplish this, we draw inspiration and perspectives from many different disciplines, as shown in figure 4. From physiology we seek to understand the biological implementation of human intelligence; from mathematics, the nature of information and intelligence regardless of its implementation; and from information processing, we map these insights and others into the metaphors of computer science that will most directly lead to successful computer implementation. This section describes each of these underpinnings for our road-mapping efforts.

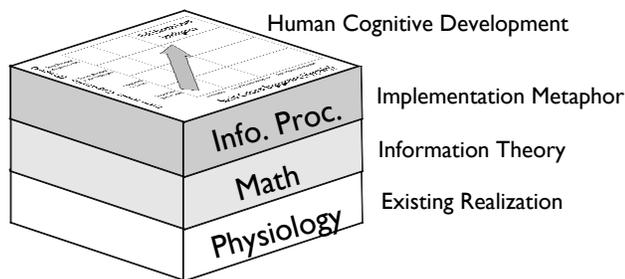


figure 4. The Substrata for the AGI Landscape

### The Physiological Perspective

Given advances in bioscience, it has become increasingly feasible to take a more holistic physiological perspective on cognitive development incorporating genetic, biochemical, and neural mechanisms, among others. A unique strength of biologically-inspired cognitive theories is their ability to account for universal aspects of development that emerge in a consistent pattern across a wide range of cultures, physical environments, and historical time-periods (e.g., motor-skill development, object perception, language acquisition, etc.). The physiological approach also plays a prominent role in explaining differences between typical and atypical patterns of development (e.g., autism, ADHD, learning disorders, disabilities, etc.). Both physiological

and information-processing perspectives favor modular accounts of cognitive development, which view the brain as divided into special-purpose input-output systems that are devoted to specific processing tasks.

### The Mathematical Perspective

This perspective is typified by the recent work of Marcus Hutter and Shane Legg, who give a formal definition of general intelligence based on the Solomonoff-Levin distribution (Legg and Hutter 2007). Put very roughly, they define intelligence as the average reward-achieving capability of a system, calculated by averaging over all possible reward-summable environments, where each environment is weighted in such a way that more compactly describable programs have larger weights. Variants of this definition have been proposed to take into account the way that intelligence may be biased to particular sorts of environments, and the fact that not all intelligent systems are explicitly reward-seeking (Goertzel 2010).

While this notion of intelligence as compression is useful, even if a mathematical definition of general intelligence were fully agreed upon, this wouldn't address the "human-level" part of AGI. Human intelligence is neither completely general in the sense of a theoretical AGI like AIXI, Hutter's "optimally intelligent agent" (Hutter 2004), nor is it highly specialized in the sense of current AI software. It has strong general capability, yet is biased toward the class of environments in which human intelligence develops – a class of environments whose detailed formalization remains largely an unsolved problem.

### The Information Processing Perspective

Much of modern cognitive science uses the computer as a metaphor for the mind. This perspective differs from the system-theoretic ideas of Piaget and Vygotsky, but at the same time provides a more direct mapping to the target implementation for AGI systems.

According to the information-processing perspective, cognitive development in infancy and childhood is due to changes in both "hardware" (neural maturation, e.g., synaptogenesis, neural pruning, etc.) and "software" (e.g., acquisition of strategies for acquiring and processing information). Rather than advocating a stage-based approach, this perspective often highlights processes of change that are gradual or continuous. For example, short-term memory – typically measured by the ability to remember a string of random letters or digits – improves linearly during early childhood. (Dempster 1981)

### Challenge: Finding Tasks and Environments

The next challenge we address is that of defining appropriate tasks and environments for assessing progress in AGI where all of the Laird and Wray characteristics are active,

and thus all of the requirements must be confronted. The usefulness of any task or environment for this purpose is critically dependent on how well it provides a basis for comparison between alternative approaches and architectures. With this in mind, we believe that both tasks and environments must be designed or specified with some knowledge of each other. For example, consider the competition to develop an automobile capable of driving itself across a roughly specified desert course (DARPA 2004). While much of the architecture and implementation would likely be useful for the later competition for autonomous city driving (DARPA 2007), such a different environment required significant reconsideration of the tasks and sub-tasks themselves.

Tasks, therefore, require a context to be useful, and environments, including the AGI's embodiment itself, must be considered in tandem with the task definition. Using this reasoning, we decided to define scenarios that combine both tasks and their necessary environments.

Further, since we are considering multiple scenarios as well as multiple approaches and architectures, it is also important to be able to compare and connect tasks belonging to different scenarios. With this in mind, we chose to proceed by first articulating a rough heuristic list of human intelligence competencies (figure 5). As a rule of thumb, tasks may be conceived as ways to assess competencies within environments. However, contemporary cognitive science does not give us adequate guidance to formulate anything close to a complete, rigid and refined list of competencies. What we present here must be frankly classified as an intuitive approach for thinking about task generation, rather than a rigorous analytical methodology from which tasks can be derived.

## Environments and Embodiments for AGI

General intelligence in humans develops within the context of a human body, complete with many thousands of input sensors and output effectors, which is itself situated within the context of a richly reactive environment, complete with many other humans, some of which may be caregivers, teachers, collaborators or adversaries. Human perceptual input ranges from skin sensors reacting to temperature and pressure, smell and taste sensors in the nose and mouth, sound and light sensors in the ears and eyes, to internal proprioceptive sensors that provide information on gravitational direction and body joint orientation, among others. Human output effectors include the many ways we can impact the physical states of our environment, like producing body movements via muscles, sound waves using body motions or producing vibrations via vocal chords, for example. Effectors like the muscles controlling gaze direction and focus of the eyes also directly impact environmental sensing.

One of the distinctive challenges of achieving AGI is thus how to account for a comparably rich sensory/motor experience. The history of AI is replete with various attempts at robotic embodiment, from Gray Walter's Tortoise (Walter 1953) through Honda's ASIMO. Virtual embodiment has also been utilized, exploiting software systems that simulate bodies in virtual environments, such as Steve Grand's "Creatures." (Grand 2001)

There are many challenges in determining an appropriate level of body sophistication and environmental richness in designing either physical or virtual bodies and environments. Regrettably, nearly every AI or AGI project that has included situated embodiment has done so in some unique, incompatible and typically non-repeatable form. A notable

figure 5. Some of the Important Competency Areas Associated with Human-Level General Intelligence

|                   |                                    |                      |                           |  |                            |
|-------------------|------------------------------------|----------------------|---------------------------|--|----------------------------|
| <b>Perception</b> | <b>Memory</b>                      | <b>Attention</b>     | <b>Social Interaction</b> | <b>Planning</b>  | <b>Motivation</b>          |
| Vision            | Working                            | Visual               | Communication             | Tactical   | Sub-goal creation          |
| Smell             | Episodic                           | Auditory             | Appropriateness           | Strategic  | Affect-based               |
| Touch             | Implicit                           | Social               | Social Inference          | Physical   | Deferred Gratification     |
| Taste             | Semantic                           | Behavioral           | Cooperation               | Social   | Altruism                   |
| Audition          | Procedural                         |                      | Competition               |  |                            |
| Cross-modal       |                                    | <b>Communication</b> | Relationships             |  | <b>Modeling Self/Other</b> |
| Proprioceptive    | <b>Reasoning</b>                   | Gestural             |                           | <b>Emotion</b>   | Self-awareness             |
|                   | Induction                          | Verbal               | <b>Learning</b>           | Perceived  | Other-awareness            |
| <b>Actuation</b>  | Deduction                          | Musical              | Imitation                 | Expressed  | Relationships              |
| Physical Skills   | Abduction                          | Pictorial            | Reinforcement             | Control  | Self-control               |
| Tool Use          | Physical                           | Diagrammatical       | Dialogical                | Understanding  | Theory of Mind             |
| Navigation        | Causal                             | Language Acquisition | Media-oriented            | Sympathy   | Sympathy                   |
| Proprioceptive    | Associational                      | Cross-modal          | Experimentation           | Empathy  | Empathy                    |
|                   | <b>Building/creation</b>           |                      |                           | <b>Quantitative</b>  |                            |
|                   | Physical construction with objects |                      |                           | Counting observed entities                                 |                            |
|                   | Formation of novel concepts        |                      |                           | Grounded small number arithmetic                           |                            |
|                   | Verbal invention                   |                      |                           | Comparison of quantitative properties of observed entities |                            |
|                   | Social organization                |                      |                           | Measurement using simple tools                             |                            |

exception to this have been various leagues of the Robocup competition (Kitano 1997), where a commercially available hardware robotic platform such as the Sony AIBO or Aldebaran Nao is specified, with contestants only allowed to customize the software of the robotic players. Unfortunately, the limited nature of the Robocup task – a modified game of soccer, is too narrow to be suitable as a benchmark for AGI.

### **The Breadth of Human Competencies**

Regardless of the particular environment or embodiment chosen, there are, intuitively speaking, certain competencies that a system must possess to be considered a full-fledged AGI. For instance, even if an AI system could answer questions involving previously known entities and concepts vastly more intelligently than IBM’s WATSON, we would be reluctant to classify it as human-level if it could not create new concepts when the situation called for it.

While articulation of a precise list of the competencies characterizing human level general intelligence is beyond the scope of current science (and definitely beyond the scope of this paper), it is nevertheless useful to explore the space of competencies, with a goal of providing heuristic guidance for the generation of tasks for AGI systems acting in environments. Figure 5 lists a broad set of competencies that we explored at the road-mapping workshop: 14 high level competency areas, with a few key competency sub-areas corresponding to each. A thorough discussion of any one of these sub-areas would involve a lengthy review paper and dozens of references, but for the purposes of this paper an evocative list will suffice.

We consider it important to think about competencies in this manner -- because otherwise it is too easy to pair a testing environment with an overly limited set of tasks biased to the limitations of that environment. What we are advocating is not any particular competency list, but rather the approach of exploring a diverse range of competency areas, and then generating tasks that evaluate the manifestation of one or more articulated competency areas within specified environments.

Some of the competencies listed may appear intuitively simpler and “earlier-stage” than others, at least by reference to human cognitive development. The developmental theories reviewed earlier in this paper each contain their own hypotheses regarding the order in which developing humans achieve various competencies. However, after careful consideration we have concluded that the ordering of competencies isn’t necessary for an AGI capability roadmap. In many cases, different existing AGI approaches pursue the same sets of competencies in radically different orders, so imposing any particular ordering on competencies would bias the road-mapping effort toward particular

AGI approaches, something antithetical to the collaboration and cooperation sorely needed in the present stage of AGI development.

For example, developmental robotics (Lungarella 2003) and deep learning (Bengio 2009) AGI approaches often assume that perceptual and motor competencies should be developed prior to linguistic and inferential ones. On the other hand, logic-focused AGI approaches are often more naturally pursued via beginning with linguistic and inferential competencies, and moving to perception and actuation in rich environments only afterwards. By dealing with scenarios, competencies and tasks, but not specifying an ordering of competencies or tasks, we are able to present a capability roadmap that is friendly to existing approaches and open to new approaches as they develop.

### **Supporting Diverse Concurrent Research Efforts**

Balancing the AGI community’s need for collaborative progress while still supporting a wide range of independent research and development efforts, we must allow for concurrent activity in many regions of the landscape. This requires that many of the unsolved “lower-level” aspects of general intelligence such as visual perception and fine-motor control be “finessed” in some of our scenarios, replaced by an approximate high-level implementation or even a human surrogate, until we have sufficiently rich underlying systems as a common base for our research. It is sometimes argued that since human intelligence emerges and develops within a complex body richly endowed with innate capabilities, AGI may only be achieved by following the same path in a similar embodiment. But until we have the understanding and capability to construct such a system, our best hope for progress is to foster concurrent but ultimately collaborative work.

### **Scenarios for Assessing AGI**

In order to create a roadmap toward human-level AGI, one must begin with one or more particular scenarios. By a “scenario” we mean a combination of tasks to be performed within a specified environment, together with a set of assumptions pertaining to the way the AGI system interacts with the environment along with the inanimate entities and other agents therein. Given a scenario as a basis, one can then talk about particular subtasks, their ordering and performance evaluation.

A wide variety of scenarios may be posited as relevant to AGI; here we review seven that were identified by one or more participants in the AGI Roadmap Workshop as being particularly interesting. Each of these scenarios may be described in much more detail than we have done here; this section constitutes a high-level overview, and more details on several of the scenarios is provided in the references.

## General Video-game Learning

This scenario addresses many challenges at once: providing a standard, easily accessible situated embodiment for AGI research; a gradient of ever increasing sensory/motor requirements to support gradual development, measurement and comparison of AGI capabilities; a wide range of differing environments to test generality of learning ability; and a compelling demonstration of an AGI's capability or lack thereof, easily evaluated by non-specialists, based on widely common human experiences playing various video games.

The goal of this scenario would not be human-level performance of any single video game, but the ability to learn and succeed at a wide range of video games, including new games unknown to the AGI developers before the competition. The contestant system would be limited to a uniform sensory/motor interface to the game, such as video and audio output and controller input, and would be blocked from any access to the internal programming or states of the game implementation. To provide for motivation and performance feedback during the game, the normal scoring output would be mapped to a uniform hedonic (pain/pleasure) interface for the contestant. Contestants would have to learn the nature of the game through experimentation and observation, by manipulating game controls and observing the results. The scores achieved against a pre-programmed opponent or storyline would provide a standard measure of achievement along with the time taken to learn and win or advance in each kind of game.

The range of video games used for testing in this scenario could be open ended in both simplicity and sophistication (History 2011). Early games like *Pong* might itself prove too challenging a starting point, so even simpler games may be selected or developed. Since success at most video games would require some level of visual intelligence, General Video-game Learning (GVL) would also provide a good test of computer vision techniques, ranging from relatively simple 2D object identification and tracking in *Pong* to full 3D perception and recognition in a game like *Tomb Raider* or *Half-Life* at the highest levels of performance. Various genres of video games such as the early side-scrolling games (e.g., *Super Mario Brothers*) to first person shooters (e.g., *Doom*, *Half-Life*) to flight simulations (e.g., *Microsoft Flight Simulator*, *Star Wars X-Wing*) provide rich niches where different researchers could focus and excel, while the common interface would still allow application of learned skills to other genres.

Among other things, in order to effectively play many videos games a notable degree of strategic thinking must be demonstrated, such as the ability to map situations to actions while considering not just the short-term but also the longer-term implications of choosing an action. Such capability, often associated with the notion of the credit assign-

ment problem, is one that remains to be demonstrated in scalable way. GVL provides a effective platform for such demonstration.

This scenario has the added benefit of supporting the growth of a research community focused on the specification and development of appropriate tests for AGI. Such a community would create the video game tests, in collaboration with the original game vendors in many cases, and both administer the tests and report the results via an internet forum similar to the SPEC-style system performance reporting sites (SPEC 2011).

## Preschool Learning

In the spirit of the popular book “All I Really Need to Know I Learned in Kindergarten” (Fulghum 1989), it is appealing to consider early childhood education such as kindergarten or preschool as inspiration for scenarios for teaching and testing AGI systems. The details of this scenario are fleshed out in (Goertzel and Bugaj 2009).

This scenario has two obvious variants: a physical preschool-like setting involving a robot, and a virtual-world preschool involving a virtual agent. The goal in such a scenarios is not to precisely imitate human child behavior, but rather demonstrate a robot or virtual agent qualitatively displaying similar cognitive behaviors to a young human child. This idea has a long and venerable history in the AI field – Alan Turing's original 1950 paper on AI, where he proposed the famous “Turing Test”, contains the suggestion that “Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's?” (Turing 1950)

This “childlike cognition” based approach seems promising for many reasons, including its integrative nature: what a young child does involves a combination of perception, actuation, linguistic and pictorial communication, social interaction, conceptual problem solving and creative imagination. Human intelligence develops in response to the demands of richly interactive environments, and a preschool is specifically designed to be just such an environment with the capability to stimulate diverse mental growth. The richness of the preschool environment suggests that significant value is added by the robotics based approach; but much could also potentially be achieved by stretching the boundaries of current virtual world technology.

Another advantage of focusing on childlike cognition is the wide variety of instruments for measuring child intelligence developed by child psychologists. In a preschool context, the AGI system can be presented with variants of tasks typically used to measure the intelligence of young human children.

It doesn't currently make sense to outfit a virtual or robot preschool as a precise imitation of a human preschool – this

would be inappropriate since a contemporary robotic or virtual body is rather differently capable than that of a young human child. The aim in constructing an AGI preschool environment should be to emulate the basic diversity and educational character of a typical human preschool.

To imitate the general character of a human preschool, several “learning centers” should be available in a virtual or robot preschool. The precise architecture will be adapted via experience but initial learning centers might include:

- The Blocks Center: a table with blocks of various shapes and sizes on it
- The Language Center: a circle of chairs, intended for people to sit around and talk with the AGI
- The Manipulatives Center, with a variety of different objects of different shapes and sizes, intended to teach a wide range of visual and motor skills
- The Ball Play Center: where balls are kept in chests and there is space for the AGI to kick or throw the balls around alone or with others
- The Dramatics Center where the AGI can both observe and enact various movements and roles

## Reading Comprehension

The next scenario is closely related to the previous one, but doesn’t require embodiment in a virtual world, and makes a special commitment as to the type of curriculum involved. In this scenario, an aspiring AGI should work through the grade school reading curriculum, take and pass the assessments normally used to assess the progress of human children. This obviously requires understanding a natural language (NL) text and being able to answer questions about it. However, it also requires some not so obvious abilities.

Very early readers are usually picture books that tightly integrate the pictures with the text. In some, the story is mostly conveyed through the pictures. In order to understand the story, the pictures must be understood as well as the NL text. This requires recognizing the characters and what the characters are doing. Reference resolution is required between characters and events mentioned in the text and illustrated in the pictures. The actions that the characters are performing must be recognized from “snapshot” poses, unlike the more usual action recognition from a sequence of frames taken from a video. The next stage of readers are “early chapter books,” which use pictures to expand on the text. Although the story is now mainly advanced through the text, reference resolution with the pictures is still important for understanding.

Instructors of reading recognize “four roles of a reader as: meaning maker, code breaker, text user, and text critic... meaning makers read to understand... comprehension questions [explore] literal comprehension, inferential comprehension, and critical thinking” (Sundance 2001). Code

breakers translate written text to speech and vice versa. Text users identify whether the text is fiction or nonfiction. “If the book is factual, they focus on reading for information. If a text is fiction, they read to understand the plot, characters, and message of the story... Text critics evaluate the author’s purpose and the author’s decisions about how the information is presented ... check for social and cultural fairness ... look for misinformation ... and think about their own response to the book and whether the book is the best it might be” (Sundance 2001).

The roles of meaning maker, code breaker, and text user are mostly, though not entirely, familiar to people working in Natural Language Understanding. However, the role of text critic is new, requiring meta-level reasoning about the quality of the text, its author’s intentions, the relation of the text to the reader’s society and culture, and the reader’s own reaction to the text, rather than merely reasoning about the information in the text.

This scenario fulfills most of the Laird and Wray criteria handily, and a few less obviously. The AGI’s environment in this scenario is not dynamic in a direct sense (C2), but the AGI does have to reason about a dynamic environment to fulfill the tasks. In a sense, the tasks involved are fixed rather than novel (C5), but they are novel to the AGI as it proceeds through them. Other agents impact task performance (C4) if group exercises are involved in the curriculum (which indeed is sometime the case). Many of the specific abilities needed for this scenario are discussed in the next scenario.

## Story/Scene Comprehension

Focused on scene and story comprehension, this scenario shares aspects of the preschool and reading comprehension scenarios. Like the reading scenario, it focuses on a subset of child-like performance – but a different subset, involving a broader variety of engagements with the world. “Scene comprehension” here does not mean only illustrations, but real-world scenes, which can be presented at different granularities, media and complexities (cartoons, movies, or theatrical performances for instance). This approach differs from the reading scenario in that it more directly provides a dynamic environment. If group exercises are included then all the Laird and Wray criteria are fulfilled in a direct and obvious way.

The story/scene comprehension scenario focuses on the relationship between perception, mental representation and natural language. Here, the system might be presented with a story which it needs to analyze semantically and then re-present in a different form (for instance, as a movie sequence, or another kind of re-telling). In the other direction, the system is shown a movie or cartoon sequence, which it needs to comprehend and re-present using language. This approach lends itself to a variety of standardizable test sce-

narios, which allow the direct comparison of competing AGI architectures with each other, as well as with child performance using identical input.

Note that these tests ignore the interaction of the system and the environment, and the particular needs of the system itself. It is often argued that general intelligence might only be attained by a motivated, interacting system, but this scenario would allow the comparison of motivated, interactive architectures with less comprehensive processing oriented architectures.

AGI system capabilities in story and scene comprehension are probably not independent but related, since both tasks focus on modeling the relationships between perception, language and thought, which are not unidirectional. Further tasks might focus on the interplay between these areas, by combining expression and comprehension. This could either be done in a fully interactive manner (i.e., by designing a system that has to respond to a teacher), or “self interactive” (two agents learn to communicate with each other, while sharing an environment). For an example design of such a self-interactive scenario, see (Steels 2010).

### **School Learning**

The “virtual school student” scenario continues the virtual preschool scenario discussed above but is focused on higher cognitive abilities assuming that, if necessary, lower-level skills will be finessed. In particular, it is assumed that all interfaces with the agent are implemented at a symbolic level: the agent is not required to process a video stream, to recognize speech and gestures, to balance its body and avoid obstacles while moving in space, etc. All these can be added later to enhance the challenge, but a key feature of this scenario is that they can also be finessed, allowing for early exploration of such capabilities without waiting for other system developments. On the other hand, it is critically important in the scenario for the agent to make academic progress at a human student level, to understand human minds, and to understand and use classroom-related social relations in the environment in which it is embedded.

In this scenario, the agent is embedded in a real high school classroom by means of a virtual-reality-based interface. The agent lives in a symbolic virtual world that is continuously displayed on a big screen in the classroom. The virtual world includes a virtual classroom represented at a symbolic (object) level, including the human instructor and human students represented by simplistic avatars. The agent itself is represented by an avatar in this virtual classroom. The symbolic virtual world is “synchronized” with the real physical world with the assistance of intelligent monitoring and recording equipment performing scene analysis, speech recognition, language comprehension, gesture recognition, etc. (if necessary, some or all of these

functions will be performed by hidden human personnel running the test; students should not be aware of their existence). The study material, including the textbook and other curriculum materials available to each student, will be encoded electronically and made available to the agent at a symbolic level.

The second part of the challenge means that the agent will be evaluated not only on its own learning and problem solving performance, but also its approach to problem solving, based on its interactions with students and with the instructor. Here the general metrics for self-regulated learning will be used (Winne and Perry 2000). In addition, social performance of the agent can be evaluated based on surveys of students and using standard psychological metrics. Another potentially important measure will be the effect of the agent presence in the classroom on human student learning.

### **The Wozniak Test**

In an interview a few years ago, Steve Wozniak of Apple Computer fame expressed doubt that there would ever be a robot that could walk into an unfamiliar house and make a cup of coffee (Wozniak and Moon 2007). We feel that this task is demanding enough to stand as a “Turing Test” equivalent for embodied AGI. Note that the Wozniak Test is a single, special case of Nils Nilsson’s general “Employment Test” for Human-Level AI (Nilsson 2005).

A robot is placed at the door of a typical house or apartment. It must find a doorbell or knocker, or simply knock on the door. When the door is answered, it must explain itself to the householder and enter once it has been invited in. (We will assume that the householder has agreed to allow the test in her house, but is otherwise completely unconnected with the team doing the experiment, and indeed has no special knowledge of AI or robotics at all.) The robot must enter the house, find the kitchen, locate local coffee-making supplies and equipment, make coffee to the householder’s taste, and serve it in some other room. It is allowed, indeed required by some of the specifics, for the robot to ask questions of the householder, but it may not be physically assisted in any way.

The current state of the art in robotics falls short of this capability in a number of ways. The robot will need to use vision to navigate, identify objects, possibly identify gestures (“the coffee’s in that cabinet over there”), and to coordinate complex manipulations. Manipulation and physical modeling in a tight feedback learning loop may be necessary, for example, to pour coffee from an unfamiliar pot into an unfamiliar cup. Speech recognition, natural language understanding and generation will be necessary. Planning must be done at a host of levels ranging from manipulator paths to coffee brewing sequences.

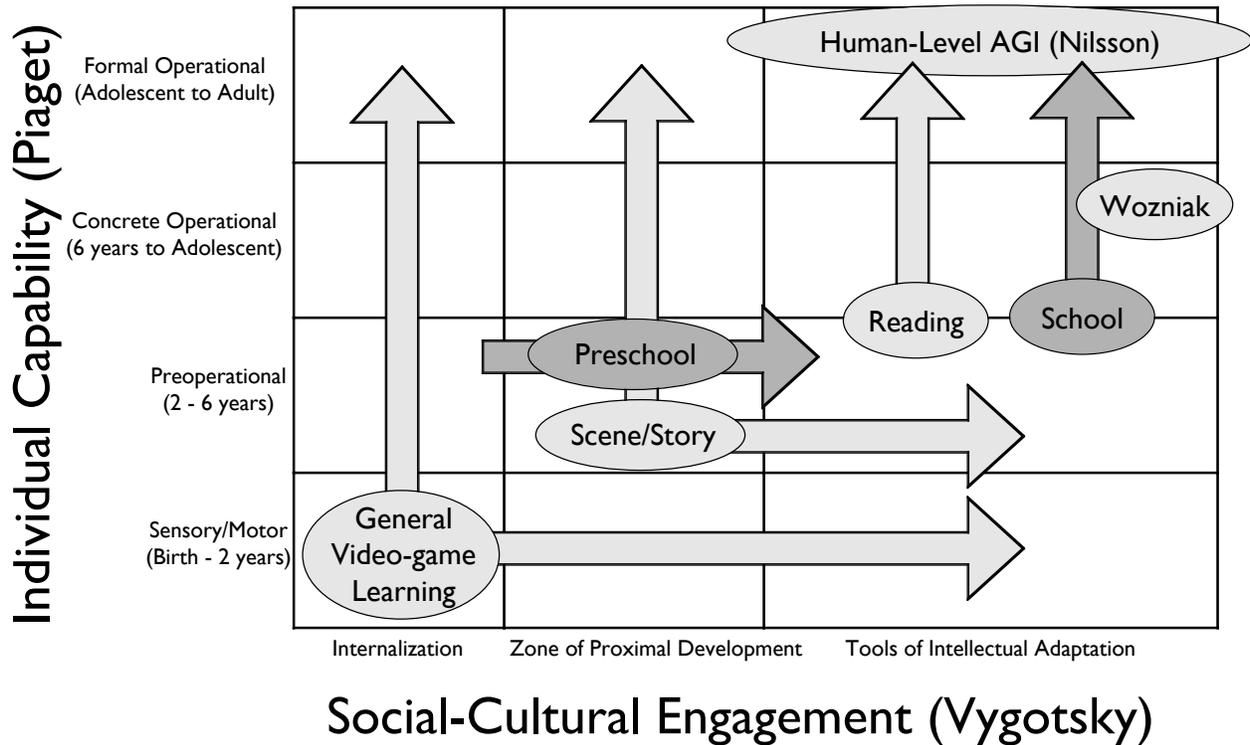


figure 6. Scenario Milestones on the AGI Landscape

The major AGI advance for a coffee-making robot is that all of these capabilities must be coordinated and used appropriately and coherently in aid of the overall goal. The usual set-up, task definition, and so forth are gone from standard narrow AI formulations of problems in all these areas; the robot has to find the problems as well as to solve them. That makes coffee-making in an unfamiliar house a strenuous test of a system’s adaptivity and ability to deploy common sense.

Although standard shortcuts might be used, such as having a database of every manufactured coffeemaker built in, it would be prohibitive to have the actual manipulation sequences for each one pre-programmed, especially given the variability in workspace geometry, dispensers and containers of coffee grounds, and so forth. Transfer learning, generalization, reasoning by analogy, and in particular learning from example and practice are almost certain to be necessary for the system to be practical.

Coffee-making is a task that most 10-year-old humans can do reliably with a modicum of experience. A week’s worth of being shown and practicing coffee making in a variety of homes with a variety of methods would provide the grounding for enough generality that a 10 year old could make coffee in the vast majority of homes in a Wozniak Test. Another advantage to this test is it would be extremely difficult to “game” or cheat, since the only reasonably economical way to approach the task would be to build general learning skills and have a robot that is capa-

ble of learning not only to make coffee but any similar domestic chore.

### An Initial Road Map for Achieving AGI

Having set forth these seven scenarios, we can now proceed to array them across the previously defined AGI landscape (figure 3). Figure 6 shows the result of this mapping, with some scenarios like General Video-game Learning and Scene/Story Comprehension spanning large areas of the landscape as their content increases in complexity and cognitive challenge. Just as we consider the Laird and Wray requirements as a practical starting point and our scenarios as initial representatives of a much larger, more representative collection, we hope this initial road map is enhanced and corrected as our collective experience in AGI grows.

### Remaining Whitespace

In our view, the suggested scenarios represent a significant step in populating the AGI Landscape, but many alternative scenarios are conceivable, and might even cover areas that we have incompletely addressed or not mentioned at all.

Among the areas that might need more attention, are:

- Aesthetic appreciation and performance: Composition, literature, artistic portrayal, dance, etc. Examining the

appreciation of aesthetics is bound to increase our understanding of the motivational system, of mental representation, and of meta-cognition. Aesthetic performance adds manipulation, creation and adaptation of physical and imagined objects, Theory of Mind, aspects of sociality and many things more.

- Structured social interaction: Goal adoption, collaboration, competition and exploitation, negotiation, discourse and joint decision making, group organization, leadership and so on. Scenarios in the social domain will require representations and assessment of social setups, the mental states of individuals and possible ramifications for the agent's own goals.
- Skills that require high cognitive function and integration, for instance, complex rescue scenarios, physical shopping, and skilled human assistance. Each of these task areas would require complex coordination between interaction, exploration, evaluation and planning.

Thus, we encourage and invite suggestions for additional scenarios, as long as these are firmly rooted in an AGI approach, i.e., each scenario should foster our understanding of general, human-like intelligence, and not simply provide narrow engineering solutions to a limited task.

## **From Scenarios to Tasks, Metrics and Challenges**

The articulation of scenarios and competency areas, as we have presented here, makes it possible to articulate specific tasks for assessing progress toward AGI in a principled manner. For each of the scenarios reviewed above (or other analogous scenarios), one may articulate a specific task-set, where each task addresses one or more of the competency areas in the context of the scenario. To constitute a reasonable approximation of an AGI test suite or overall “AGI Challenge,” the total set of tasks for a scenario must cover all the competency areas. Each task must also be associated with some particular performance metric -- quantitative wherever possible, but perhaps qualitative in some cases depending on the nature of the task.

The obvious risk of an AGI evaluation approach based on a long list of tasks is that it is susceptible to solution by the so-called ‘big switch statement’ approach, in which separate narrowly-specialized programs corresponding to the individual tasks are combined together in a simplistic harness. Some may believe that human intelligence is well approximated by a big switch statement of this sort, but we take a different perspective which doesn’t rule out the pos-

sibility, but also doesn’t focus on or encourage this approach.

In addition, the competency areas alluded to above include many that focus on learning and generalization. So even if a “big switch statement” approach is used to handle tasks based on these requirements, the switch will be choosing between a large number of specialized programs that in themselves could handle a great deal of learning.

### **Example Tasks**

For sake of brief illustration, table 1 roughly describes a handful of example tasks and task families corresponding to a few of the scenarios described above. Much of the work of fleshing out the initial roadmap described here will involve systematically creating a substantial list of tasks corresponding to the scenarios chosen, using an exploratory enumeration of human general intelligence competencies as a guide (perhaps a refined version of the list in figure 5). Each task will then need to be associated with quantitative performance metrics.

Some of the examples in table 1 are highly specific, others are broader task families. Our feeling is that it will generally be better to identify moderately broad task families rather than highly specific tasks, in order to encourage the creation of systems with more robust and flexible functionality.

### **Multidimensional Challenges**

What then is the right sort of challenge to present to a would-be AGI system? We suggest the following: to address a few dozen closely interrelated, moderately broadly-defined tasks in an environment drawn from a moderately broad class (e.g. an arbitrary preschool or textbook rather than specific ones determined in advance). If the environment is suitably rich and the tasks are drawn to reflect the spectrum of human general intelligence competencies, then this sort of challenge will motivate the development of genuine human-level AGI systems.

This sort of challenge is not nearly as well specified as a chess contest, RoboCup or the DARPA Grand Challenge, but we feel the complexity and heterogeneity here is directly reflective of the complexity and heterogeneity of human general intelligence. Even if there are simple core principles underlying human-level general intelligence, as some have posited, nevertheless the real-world manifestation of these principles - which is what we can measure - is highly complex and involves multiple interrelated competencies defined via interaction with a rich, dynamic world.

### **Challenges and Competitions**

Finally, we emphasize that our main goal here is to propose a capability roadmap for the AGI research community, not a competition or contest. Competitions can be valuable,

| Scenario               | Competency Area         | Sub-area             | Example Task or Task Family   |
|------------------------|-------------------------|----------------------|---|
| Virtual Preschool      | Learning                | Dialogical           | Learn to build a particular structure of blocks (e.g., a pyramid) faster based on a combination of imitation, reinforcement and verbal instruction, than by imitation and reinforcement without verbal instruction  |
| Virtual Preschool      | Modeling Self and Other | Theory of Mind       | While Sam is in the room, Ben puts the red ball in the red box. Then Sam leaves and Ben moves the red ball to the blue box. Sam returns and Ben asks him where the red ball is. The agent is asked where Sam thinks the ball is.  |
| Virtual School Student | Learning                | Media-oriented       | Starting from initially available basic concepts (a number, a variable, a function), demonstrate academic progress in learning how to solve problems from the textbook using techniques described in the same textbook. The agent should move step by step, from simple to advanced problems, from one domain to another. |
| Robot Preschool        | Actuation               | Proprioception       | The teacher moves the robot's body into a certain configuration. The robot is asked to restore its body to an ordinary standing position, and then repeat the configuration that the teacher moved it into.   |
| Robot Preschool        | Memory                  | Episodic             | Ask the robot about events that occurred at times when it received a particularly large, or particularly small, reward for its actions; it should be able to answer simple questions about these significant events, with significantly more accuracy than about events occurring at random times                         |
| Wozniak Test           | Communication           | Gestural             | The robot will be directed to the kitchen. It must understand gestures indicating that it should follow an indicated path, or know how to follow its guide, and know when either is appropriate.  |
| Wozniak Test           | Actuation               | Navigation           | The robot must complete its task without colliding people, walls, furniture, or pets.   |
| Wozniak Test           | Social Interaction      | Appropriate behavior | The robot must recognize and appropriately respond to the situation where it has knocked on the wrong door, the householder is not home, or the robot is not welcome to enter the house.  |
| Wozniak Test           | Reasoning               | Physical             | Be able to use incomplete or alternative tools and equipment. For instance, a drip pot may be used without its top, but not a percolator. This may require physical simulation, based on an understanding of naive physics.   |
| Wozniak Test           | Reasoning               | Induction            | On the other hand, the above-mentioned knowledge about drip pots and percolators may be gathered via inductive reasoning based on observations in multiple relevant situations.   |

Table 1. Mapping Human General Intelligence Competencies to Scenario Tasks and Task Families

and we would be happy to see one or more developed along the lines of the capability roadmap we have described. However, we doubt any successful competition could ever cover the full extent of a roadmap such as this. One reason for this is that a competition is very likely to involve concrete decisions regarding the prioritization or temporal ordering of tasks, which will inevitably result in the competition having widely differential appeal to AGI researchers depending on their chosen scientific approach.

The primary value to be derived via wide adoption of a capability roadmap of this nature is the concentration of multiple researchers pursuing diverse approaches on common scenarios and tasks -- which allows them to share

ideas and learn from each other much more than would be the case otherwise. There is much to be gained via multiple researchers pursuing the same concrete challenges in parallel, with or without the presence of an explicit judged competition.

### Call For Collaboration

We have explained our vision for the creation of a roadmap to human-level Artificial General Intelligence, selected goals and starting points, established a landscape, populated it with a number of diverse scenarios and fleshed out some of the details. What's next is to put more meat on the

bones: describe the scenarios in more detail, refine the list of specific competency areas, and then create tasks and metrics along the lines outlined above. Once this is done, we will have a capability roadmap clearly articulating some plausible pathways to AGI; and we will have the basis for creating holistic, complex challenges that can meaningfully assess progress on the AGI problem.

Such a roadmap does not give a highly rigorous, objective way of assessing the percentage of progress toward the end-goal of AGI. However, it gives a much better sense of progress than one would have otherwise. For instance, if an AGI system performed well on diverse metrics corresponding to tasks assessing 50% of the competency areas listed above, in several of the above scenarios, the creators of this system would seem justified in claiming to have made very substantial progress toward AGI. If the number were 90% rather than 50%, one would seem justified in claiming to be "almost there." If the number were 25%, this would give one a reasonable claim to "interesting AGI progress." This kind of qualitative assessment of progress is not the most one could hope for, but it is better than the progress indications one could get without this sort of roadmap.

We have already explicitly noted many of the limitations of our treatment of this challenging topic. While there is no consensus among AGI researchers on the definition of general intelligence, we can generally agree on a pragmatic goal. The diversity of scenarios presented reflects a diversity of perspectives among AGI researchers regarding which environments and tasks best address the most critical aspects of AGI. Most likely neither the tentative list of competencies nor the Laird and Wray criteria are either necessary or sufficient. There is no obvious way to formulate a precise measure of progress toward AGI based on the competencies and scenarios provided – though one can use these to motivate potentially useful approximative measures.

But in spite of its acknowledged limitations, a capability roadmap of the sort outlined here

- allows multiple researchers following diverse approaches to compare their work in a meaningful way
- allows researchers, and other observers, to roughly assess the degree of research progress toward the end goal of human-level artificial general intelligence
- allows work on related road-mapping aspects, such as tools roadmaps and study of social implications and potential future applications, to proceed in a more structured way

With this in mind, we encourage AGI researchers to join with us in the ongoing collaborative construction of a more detailed AGI roadmap along the lines suggested here.

## References

- Arel, I. 2009. Working toward pragmatic convergence: AGI axioms and a unified roadmap. In Workshop on the Future of AGI (as part of the 2nd Conference on Artificial General Intelligence, AGI-09).
- Arel, I., and Livingston, S. 2009. Beyond the Turing test. *IEEE Computer*, 42(3):90–91.
- Bengio, Y. 2009. Learning Deep Architectures for AI, *Foundations and Trends in Machine Learning*, 2, 1-127.
- Campbell, M.; Hoane, A. J. Jr.; and Hsu, F. 2002. Deep Blue. *Artif. Intell.* 134, 1-2 (January), 57-83.
- DARPA 2004. DARPA Grand Challenge rulebook. Online. [http://www.darpa.mil/grandchallenge05/Rules\\_8oct04.pdf](http://www.darpa.mil/grandchallenge05/Rules_8oct04.pdf)
- DARPA 2007. DARPA Urban Challenge Rules, 2007. Online. <http://www.darpa.mil/grandchallenge/rules.asp>
- Dempster, F. N. 1981. Memory span: Sources of individual and developmental differences. *Psychological Bulletin*, 89, 63-100.
- Ferrucci, D. 2010. Build Watson: an overview of DeepQA for the Jeopardy! challenge. In Proceedings of the 19th international conference on Parallel architectures and compilation techniques (PACT '10). ACM, New York, NY, USA, 1-2.
- Fulghum, R. 1989. All I Really Need To Know I Learned In Kindergarten. Ivy Books.
- Gardner, H. 1999. Intelligence reframed: Multiple intelligences for the 21st century. Basic.
- Goertzel, B. 2010. Toward a formal definition of real-world general intelligence. In Proceedings of the Third Conference on Artificial General Intelligence (AGI-10). Atlantis Press.
- Goertzel, B., and Bugaj, S. V. 2009. AGI preschool. In Proceedings of the Second Conference on Artificial General Intelligence (AGI09). Atlantis Press.
- Grand, S. 2001. *Creation : life and how to make it*. London: Phoenix.
- Gregory, R. 1996. *Psychological testing: History, principles, and applications*. Allyn and Bacon.
- History of Video Games 2011. Wikipedia entry. [http://en.wikipedia.org/wiki/History\\_of\\_video\\_games](http://en.wikipedia.org/wiki/History_of_video_games)
- Hutter, M. 2004. *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability*. Springer, Berlin.
- Kitano, H. 1997. RoboCup: The Robot World Cup Initiative in Proc. of The First International Conference on Autonomous Agent (Agents-97), Marina del Rey, The ACM Press
- Laird, J., and Wray, R. 2010. Cognitive architecture requirements for achieving AGI. In Proceedings of the Third Conference on Artificial General Intelligence (AGI-10). Atlantis Press.
- Laird, J.; Wray, R.; Marinier, R.; and Langley, P. 2009. Claims and challenges in evaluating human-level intelligent systems. In Proceedings of the Second Conference on Artificial General Intelligence (AGI-09). Atlantis Press.
- Legg, S., and Hutter, M. 2007. A definition of machine intelligence. *Minds and Machines*, 17.

- Lungarella, M.; Metta, G.; Pfeifer, R.; and Sandini, G. 2003. Developmental robotics: a survey. *Connection Science*, VOL 15; PART 4, pages 151-190
- Nilsson, N. J. 2005. Human-level artificial intelligence? Be serious! *AI Magazine*, 26(4):68-75.
- Piaget, J. 1953. *The Origins of Intelligence in Children*. Routledge and Kegan Paul.
- Samsonovich, A. V.; Ascoli, G. A.; and De Jong, K. A. 2006. Human-level psychometrics for cognitive architectures. In *Proceedings of the Fifth International Conference on Development and Learning (ICDL5)*.
- SPEC 2011. Standard Performance Evaluation Corporation. <http://www.spec.org>
- Steels, L. 2010. Modeling The Formation of Language in Embodied Agents: Methods and Open Challenges, pages 223–233. Springer Verlag.
- Sundance 2001. *Teacher Guide for Just Kids*. Sundance Publishing.
- Turing, A. M. 1950. Computing machinery and intelligence. *Mind* 59 (October):433-60
- Vygotsky, L. 1986. *Thought and Language*. MIT Press.
- Walter, W. G. 1953. *The Living Brain*. New York: Norton.
- Winne, P. H., and Perry, N. E. 2000. Measuring self-regulated learning, pages 531–566. Academic Press.
- Wozniak, S., and Moon, P. 2007. Three minutes with Steve Wozniak. *PC World*, July 19.

## Author Biographical Sketches



Sam S. Adams currently works for IBM Research and was appointed one of IBM's first Distinguished Engineers in 1996. His far-ranging contributions include founding IBM's first Object Technology Practice, authoring IBM's XML technical strategy, originating the concept of Service Oriented Architecture, pioneering work in Self-Configuring and Autonomic Systems, Artificial General Intelligence, End-User Mashup Programming, Massively Parallel Manycore Programming, Petascale Analytics and Data-centric Systems. He is a founding member of the AGI community and continues to research bootstrapping human-level general intelligence via the Joshua Blue project.



Itamar Arel is an Associate Professor of Electrical Engineering and Computer Science and Director of the Machine Intelligence Laboratory at the University of Tennessee. His research focus is on high-performance machine learning architectures and algorithms, with emphasis on deep learning architectures, reinforcement learning and decision making under uncertainty. Dr. Arel is a recipient of the US Department of Energy Early Career Principal Investigator (CAREER) award in 2004. He holds a B.S., M.S and Ph.D. degrees in Electrical and Computer Engineering and an M.B.A. degree, all from Ben-Gurion University in Israel.



Rod Furlan is the Chief of R&D at Quaternix Research Inc. His work explores the intersection of artificial intelligence, high-performance computing and quantitative finance. Passionate about technology and a serial autodidact, Rod is also a Singularity University alumnus and Teaching Fellow for AI and Robotics



Professor Matthew Schlesinger received his Ph.D. in developmental psychology from the University of California at Berkeley, with postdoctoral research in robotics and machine learning. He is currently an associate professor of psychology at Southern Illinois University, Carbondale, where he is a member of the Brain and Cognitive Sciences program and directs the Vision Lab. Schlesinger serves as an Associate Editor of “IEEE Transactions on Autonomous Mental Development,” in addition to co-chairing the TAMM-IEEE Technical Committee. His work integrates a number of approaches that span psychology and machine learning, and includes experimental research with human infants, children and adults, modeling of perceptual and cognitive processes in embodied agents, as well as neuroimaging.



Joscha Bach studied Computer Science and Philosophy at Berlin's Humboldt University, and Waikato University (NZ). He received a PhD in Cognitive Science from the University of Osnabrück. His main research interest concerns the design of cognitive architectures, especially with respect to representation, motivation and affect. At the moment, he is a research fellow at the Berlin School of Mind and Brain, at Humboldt University of Berlin.



Ben Goertzel is Chief Scientist of financial prediction firm Aidyia Holdings, and Chairman of AI consulting firm Novamente LLC and bioinformatics firm Biomind LLC. He is also leader of the OpenCog open-source artificial general intelligence project, General Chair of the AGI conference series, and Vice Chair of futurist organization Humanity+. His major research foci are embodied artificial general intelligence, and applications of AI to genomics.



J. Storrs Hall, PhD., is an independent scientist and author. His recent book "Beyond AI: Creating the Conscience of the Machine" (Prometheus, 2007), was the first full-length treatment of machine ethics. He is active in the machine ethics and AGI (artificial general intelligence) research communities, serving on the program committee for the international AGI conference series.



Stuart C. Shapiro is Professor of Computer Science and Engineering and Affiliated Professor of Linguistics and of Philosophy at the University at Buffalo, The State University of New York. He received an SB in mathematics from the Massachusetts Institute of Technology in 1966, and an MS and PhD in computer sciences from the University of Wisconsin - Madison in 1968 and 1971. His primary research interests are in knowledge representation and reasoning, especially in support of natural language competence and cognitive robotics. He is a past Chair of ACM/SIGART; past President of Principles of Knowledge Representation and Reasoning, Incorporated; a Life Senior Member of the IEEE; an ACM Distinguished Scientist; and a Fellow of the AAAI.



Robert Coop is a doctoral student at the University of Tennessee at Knoxville. He is currently studying machine learning, with a focus on generalized pattern extraction methods.



Alexei V. Samsonovich is currently a Research Assistant Professor at Krasnow Institute for Advanced Study at George Mason University. He holds a M.Sc. in Theoretical Physics from the Moscow Institute of Physics and a M.Sc. and Ph.D. in Applied Mathematics from the University of Arizona. He organized the first conference on Biological Inspired Cognitive Architecture (BICA) in 2008 and has chaired the event ever since. His current research focuses on semantic cognitive mapping of natural language and biologically inspired cognitive microcircuits.



John F. Sowa spent thirty years working on research and development projects at IBM and is a cofounder of VivoMind Research, LLC. He has a BS in mathematics from MIT, an MA in applied mathematics from Harvard, and a PhD in computer science from the Vrije Universiteit Brussel. He is a fellow of the American Association for Artificial Intelligence. The language of conceptual graphs, which he designed, has been adopted as one of the three principal dialects of the ISO/IEC standard for Common Logic.



Dr. Matthias Scheutz is an Associate Professor of Computer Science at Tufts University School of Engineering. His current research interests include agent-based modeling, cognitive modeling, complex cognitive and affective robots for human-robot interaction, computational models of human language processing in mono- and bilinguals, cognitive architectures, distributed agent architectures, and interactions between affect and cognition. Matthias earned a Ph.D. in Philosophy from the University of Vienna in 1995 and a Joint Ph.D. in Cognitive Science and Computer Science from the Indiana University Bloomington in 1999.