AI-Complete, AI-Hard, or AI-Easy – Classification of Problems in AI

Roman V. Yampolskiy

Computer Engineering and Computer Science
University of Louisville
Louisville, USA
roman.yampolskiy@louisville.edu

Abstract— The paper contributes to the development of the theory of AI-Completeness by formalizing the notion of AI-Complete and AI-Hard problems. The intended goal is to provide a classification of problems in the field of General Artificial Intelligence. We prove Turing Test to be an instance of an AI-Complete problem and further show numerous AI problems to be AI-Complete or AI-Hard via polynomial time reductions. Finally, the paper suggests some directions for future work on the theory of AI-Completeness.

Keywords- AI-Complete, AI-Easy, AI-Hard, Human Oracle.

I. INTRODUCTION

Since its inception in the 1950s the field of Artificial Intelligence has produced some unparalleled accomplishments while at the same time failing to formalize the problem space it is concerned with. This paper proposes to address this shortcoming by contributing to the theory of AI-Completeness, a formalism designed to do for the field of AI what notion of NP-Completeness did for computer science in general. It is our belief that such formalization will allow for even faster progress in solving the remaining problems in humankind's conquest to build an intelligent machine.

According to the encyclopedia Wikipedia the term "AI-Complete" was proposed by Fanya Montalvo in the 1980s (Wikipedia Retrieved January 7, 2011). A somewhat general definition of the term included in the 1991 Jargon File (Raymond March 22, 1991) states:

"**AI-complete***: [MIT, Stanford, by analogy with `NP-complete'] adj. Used to describe problems or subproblems in AI, to indicate that the solution presupposes a solution to the `strong AI problem' (that is, the synthesis of a human-level intelligence). A problem that is AI-complete is, in other words, just too hard. Examples of AI-complete problems are `The Vision Problem', building a system that can see as well as a human, and `The Natural Language Problem', building a system that can understand and speak a natural language as well as a human. These may appear to be modular, but all attempts so far (1991) to solve them have foundered on*

the amount of context information and `intelligence' they seem to require."

As such, the term "AI-Complete" (or sometimes AI-Hard) has been a part of the field for many years and has been frequently brought up to express difficulty of a specific problem investigated by researchers (see (Mallery 1988; Ide and Véronis 1998; Gentry, Ramzan et al. 2005; Nejad April 2010; Bergmair December 2004; McIntire, Havig et al. July 21-23, 2009 ; Navigli and Velardi July 2005; Mueller March 1987; McIntire, McIntire et al. May 18-22, 2009; Chen, Liu et al. November 30, 2009; Mert and Dalkilic September 14-16, 2009 ; Leahu, Sengers et al. September 21 - 24, 2008; Phillips and Beveridge September 28-30,. 2009; Hendler September 2008)). This informal use further encouraged similar concepts to be developed in other areas of science: Biometric-Completeness (Phillips and Beveridge September 28-30,. 2009), ASR-Complete (Morgan, Baron et al. April 6-10, 2003). Recently numerous attempts to formalize what it means to say that a problem is "AI-Complete" have been published (Ahn, Blum et al. 2003; Demasi, Szwarcfiter et al. March 5-8, 2010; Dafna Shahaf and Amir March 26-28, 2007). Even before such formalization attempts, systems which relied on humans to solve problems which were perceived to be AI-Complete were utilized:

- **AntiCaptcha** systems use humans to break CAPTCHA security protocol(Ahn, et al. 2003; Roman V. Yampolskiy 2007a, 2007b; Roman V Yampolskiy and Govindaraju 2007; McDaniel and Yampolskiy 2011) either by directly hiring cheap workers in developing countries (Bajaj April 25, 2010) or by rewarding correctly solved CAPTCHAs with presentation of pornographic images (Vaas December 1, 2007).

- **Chinese Room** philosophical argument by John Searle shows that including a human as a part of a computational system may actually reduce its perceived capabilities such as understanding and consciousness (Searle 1980).

- **Content Development** online projects such as Encyclopedias (Wikipedia, Conservapedia), Libraries (Project Gutenberg, Video collections (YouTube) and Open Source Software (SourceForge) all rely on contributions from people for content production and quality assurance.
- **Cyphermint** a check cashing system relies on human workers to compare a snapshot of a person trying to perform a financial transaction to a picture of a person who initially enrolled with the system. Resulting accuracy outperforms any biometric system and is almost completely spoof proof (see cyphermint.com for more info).
- **Data Tagging** systems entice users into providing meta-data for images, sound or video files. A popular approach involves developing an online game which as a byproduct of participation produces a large amount of accurately labeled data (Ahn June 2006).
- **Distributed Proofreaders** employ a number of human volunteers to eliminate errors in books created by relying on Optical Character Recognition process. (see pgdp.net for more info).
- **Interactive Evolutionary Computation** algorithms use humans in place of a fitness function to make judgments regarding difficult to formalize concept such as esthetic beauty or taste (Takagi 2001).
- **Mechanical Turk** is an Amazon.com's attempt at creating Artificial Artificial Intelligence. Humans are paid varying amounts for solving problems which are believed to be beyond current abilities of AI programs (see mturk.com for more info). The general idea behind the Turk has a broad appeal and the researchers are currently attempting to bring it to the masses via the Generalized Task Markets (GTM) (Horvitz 2007; Horvitz and Paek 2007; Kapoor, Tan et al. 2008; D. Shahaf and Horvitz July 2010).
- **Spam Prevention** is easy to accomplish by having humans vote on emails they receive as spam or not. If a certain threshold is reached a particular piece of email could be said to be spam with a high degree of accuracy (Dimmock and Maddison December 2004).

Recent work has attempted to formalize the intuitive notion of AI-Completeness. In particular three such endowers are worth reviewing:

In 2003 Ahn et al. (Ahn, et al. 2003) attempted to formalize the notion of an AI-Problem and the concept of AI-Hardness in the context of computer security. An AI-Problem was defined as a triple: "$\mathcal{P} = (S, D, f)$, where S

is a set of problem instances, D is a probability distribution over the problem set S, and f : S $\rightarrow$ {0; 1}* answers the instances. Let $\delta \in 2$ (0; 1]. We require that for an $\alpha > 0$ fraction of the humans H, $Pr_{x \leftarrow D} [H(x) = f(x)] > \delta$… An AI problem $\mathcal{P}$ is said to be (δ, τ)-*solved* if there exists a program A, running in time at most τ on any input from S, such that $Pr_{x \leftarrow D, r} [A_r(x)=f(x)] \geq \delta$. (A is said to be a ($\delta, \tau$) solution to $\mathcal{P}$.) $\mathcal{P}$ is said to be a (δ, τ)-*hard AI problem* if no current program is a (δ, τ) solution to $\mathcal{P}$, and the AI community agrees it is hard to find such a solution." It is interesting to observe that the proposed definition is in terms of democratic consensus by the AI community. If researchers say the problem is hard, it must be so. Also, time to solve the problem is not taken into account. The definition simply requires that some humans be able to solve the problem (Ahn, et al. 2003).

In 2007 Shahaf and Amir (Dafna Shahaf and Amir March 26-28, 2007) have published their work on the Theory of AI-Completeness. Their paper presents the concept of the Human-Assisted Turing Machine and formalizes the notion of different Human Oracles (see Section on Human Oracles for technical details). Main contribution of the paper comes in the form of a method for classifying problems in terms of human-versus-machine effort required to find a solution. For some common problems such as Natural Language Understanding (NLU) the paper proposes a method of reductions allowing conversion from NLU to the problem of Speech Understanding via Text-To-Speech software.

In 2010 Demasi et al. (Demasi, et al. March 5-8, 2010) presented their work on problem classification for Artificial General Intelligence (AGI). The proposed framework groups the problem space into three sectors:

- **Non AGI-Bound**: problems that are of no interest to AGI researchers.
- **AGI-Bound**: problems that require human level intelligence to be solved.
- **AGI-Hard**: problems that are at least as hard as any AGI Bound problem.

The paper also formalizes the notion of Human Oracles and provides a number of definitions regarding their properties and valid operations.

II. THE THEORY OF AI-COMPLETENESS

From people with mental disabilities to geniuses human minds are cognitively diverse and it is well known that different people exhibit different mental abilities. We define a notion of a Human Oracle (HO) function capable of computing any function computable by the union of all human minds. In other words any cognitive ability of any human being is repeatable by our HO. To make our Human Oracle easier to understand we provide the following illustration of the *Human* function:
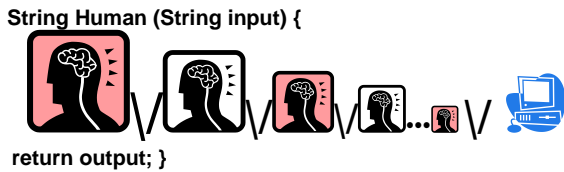
```
String Human (String input) {
```



```
return output; }
```

Figure 1.   Human oracle: Human$_{Best}$ – a union of minds.

Such a function would be easy to integrate with any modern programming language and would require that the input to the function be provided as a single string of length N and the function would return a string of length M. No specific encoding is specified for the content of strings N or M and so they could be either binary representations of data or English language phrases both being computationally equivalent. As necessary the *human* function could call regular TM functions to help in processing of data. For example, a simple computer program which would display the input string as a picture to make human comprehension easier could be executed. Humans could be assumed to be cooperating perhaps because of a reward. Alternatively, one can construct a *Human* function which instead of the union of all minds computes the average decision of all human minds on a problem encoded by the input string as the number of such minds goes to infinity. To avoid any confusion we propose naming the first HO Human$_{Best}$ and the second HO Human$_{Average}$. Problems in the AI domain tend to have a large degree of ambiguity in terms of acceptable correct answers. Depending on the problem at hand the simplistic notion of an average answer could be replaced with an aggregate answer as defined in the Wisdom of Crowds approach (Surowiecki 2004). Both functions could be formalized as Human-Assisted Turing Machines (Dafna Shahaf and Amir March 26-28, 2007).

The Human function is easy to understand and uses generalization of the Human Oracle. One can perceive it as a way to connect and exchange information with a real human sitting at a computer terminal. While easy to intuitively understand, such description is not sufficiently formal. Shahaf et al. have formalized the notion of Human Oracle as an HTM (Dafna Shahaf and Amir March 26-28, 2007). In their model a human is an oracle machine that can decide a set of languages L_i in constant time: $H \subseteq \{L_i \mid L_i \subseteq \sum^*\}$. If time complexity is taken into account answering a question might take a non-constant time: $H \subseteq \{<L_i, f_i> \mid L_i \subseteq \sum^*, f_i : \mathbb{N} \to \mathbb{N}\}$ there f_i is the time-complexity function for language L_i, meaning the human can decide if $x \in L_i$ in $f_i(|x|)$ time. In order to realistically address capabilities of individual humans a probabilistic oracle was also presented which provided correct answers with probability p: $H \subseteq \{<L_i, p_i> \mid L_i \subseteq \sum^*, 0 \leq p_i \leq 1\}$. Finally the notion of reward is introduced into the model to capture humans improved performance on "paid" tasks: $H \subseteq \{<L_i, u_i> \mid L_i \subseteq \sum^*, u_i : \mathbb{N} \to \mathbb{N}\}$

where u_i is the utility function (Dafna Shahaf and Amir March 26-28, 2007).

A.   Definitions

Definition 1: A problem C is **AI-Complete** if it has two properties:

1.   It is in the set of AI problems (Human Oracle solvable).

2.   Any AI problem can be converted into C by some polynomial time algorithm.

Definition 2: **AI-Hard:** A problem H is AI-Hard if and only if there is an AI-Complete problem C that is polynomial time Turing-reducible to H.

Definition 3: **AI-Easy:** The complexity class AI-easy is the set of problems that are solvable in polynomial time by a deterministic Turing machine with an oracle for some AI problem. In other words, a problem X is AI-easy if and only if there exists some AI problem Y such that X is polynomial-time Turing reducible to Y. This means that given an oracle for Y, there exists an algorithm that solves X in polynomial time.

Figure 2 illustrates relationship between different AI complexity classes. Right side illustrates the situation if it is ever proven that AI-problems = AI-Complete problems. Left side shows the converse.
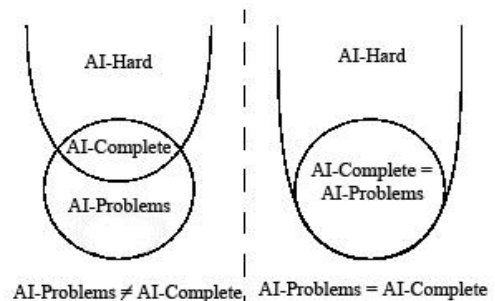

Figure 2.   Relationship between AI complexity classes.

B.   Turing Test as the First AI-Complete Problem

In this section we will show that a Turing Test (A Turing 1950) problem is AI-Complete. First we need to establish that Turing Test is indeed an AI problem (HO solvable). This trivially follows from the definition of the test itself. The test measures if a human-like performance is demonstrated by the test taker and Human Oracles are defined to produce human level performance. While both "human" and "intelligence test" are intuitively understood terms we have already shown that Human Oracles could be expressed in strictly formal terms. The Turing Test itself also could be formalized as an interactive proof (Bradford and Wollowski 1995; Shieber December 2007, July 16-20, 2006).

Second requirement for a problem to be proven to be AI-Complete is that any other AI problem should be convertible into an instance of the problem under consideration in polynomial time via Turing reduction. Therefore we need to show how any problem solvable by the Human function could be encoded as an instance of a Turing Test. For any HO-solvable problem h we have a String *input* which encodes the problem and a String *output* which encodes the solution. By taking the *input* as a question to be used in the TT and *output* as an answer to be expected while administering a TT we can see how any HO-solvable problem could be reduced in polynomial time to an instance of a Turing Test. Clearly the described process is in polynomial time and by similar algorithm any AI problem could be reduced to TT. It is even theoretically possible to construct a complete TT which utilizes all other problems solvable by HO by generating one question from each such problem.

C. *Reducing Other Problems to TT*

Having shown a first problem (Turing Test) to be AI-Complete the next step is to see if any other well-known AI-problems are also AI-complete. This is an effort similar to the work of Richard Carp who has shown some 21 problems to be NP-Complete in his 1972 paper and by doing so started a new field of Computational Complexity (Karp 1972). According to the Encyclopedia of Artificial Intelligence (Shapiro 1992) published in 1992 the following problems are all believed to be AI-Complete and so will constitute primary targets for our effort of proving formal AI-Completeness on them (Shapiro 1992):

- **Natural Language Understanding** – "Encyclopedic knowledge is required to understand natural language. Therefore, a complete Natural Language system will also be a complete Intelligent system."
- **Problem Solving** – "Since any area investigated by AI researchers may be seen as consisting of problems to be solved, all of AI may be seen as involving Problem Solving and Search".
- **Knowledge Representation and Reasoning** – "…the intended use is to use explicitly stored knowledge to produce additional explicit knowledge. This is what reasoning is. Together Knowledge representation and Reasoning can be seen to be both necessary and sufficient for producing general intelligence – it is another AI-complete area."
- **Vision or Image Understanding** – "If we take "interpreting" broadly enough, it is clear that general intelligence may be needed to do this interpretation, and that correct interpretation implies general intelligence, so this is another AI-complete area."

Now that Turing Test has been proven to be AI-Complete we have an additional way of showing other problems to be AI-Complete. We can either show that a problem is in the set of AI problems and all other AI problem can be converted into it by some polynomial time algorithm or we can reduce any instance of Turing Test problem (or any other already proven to be AI-Complete problem) to an instance of a problem we are trying to show to be AI-Complete. This second approach seems to be particularly powerful. The general heuristic of our approach is to see if all information which encodes the question which could be asked during the administering of a Turing Test could be encoded as an instance of a problem in question and likewise if any potential solution to that problem would constitute an answer to the relevant Turing Test question. Under this heuristic it is easy to see that for example Chess is not AI-Complete as only limited information can be encoded as a starting position on a standard size chess board. Not surprisingly Chess has been one of the greatest successes of AI and currently Chess playing programs dominate all human players including world champions.

Question Answering (QA) (Hirschman and Gaizauskas 2001; Salloum November 30, 2009) is a sub-problem in Natural Language Processing. Answering questions at a level of a human is something HOs are particularly good at based on their definition. Consequently QA is an AI-Problem which is one of the two requirements for showing it to be AI-Complete. Having access to an Oracle capable of solving QA allows us to solve TT via a simple reduction. For any statement S presented during administration of TT we can transform said statement into a question for the QA Oracle. The answers produced by the Oracle can be used as replies in the TT allowing the program to pass the Turing Test. It is important to note that access to the QA oracle is sufficient to pass the Turing Test only if questions are not restricted to stand alone queries, but could contain information from previous questions. Otherwise the problem is readily solvable even by today's machines such as IBM's Watson which showed a remarkable performance against human Jeopardy champions (Pepitone Retrieved on: January 13, 2011).

Speech Understanding (SU) (Anusuya and Katti 2009) is another sub-problem in Natural Language Processing. Understanding Speech at a level of a human is something HOs are particularly good at based on their definition. Consequently SU is an AI-Problem which is one of the two requirements for showing it to be AI-Complete. Having access to an Oracle capable of solving SU allows us to solve QA via a simple reduction. We can reduce QA to SU by utilizing any Text-to-Speech software (Taylor and Black 1999; Chan 2003) which is both fast and accurate. This reduction effectively transforms written questions into the spoken ones making it possible to solve every instance of QA by referring to the SU oracle.

D. Other Probably AI-Complete Problems

Figure 3 shows the relationship via reductions between problems shown to be AI-Complete in this paper. We hope that our work will challenge the AI community to prove other important problems as either belonging or not belonging to that class. While the following problems have not been explicitly shown to be AI-Complete, they are strong candidates for such classification and are also problems of great practical importance making their classification a worthy endower. If a problem has been explicitly conjectured to be AI-Complete in a published paper we include a source of such speculation: Dreaming (Salloum November 30, 2009), Commonsense Planning (Dafna Shahaf and Amir March 26-28, 2007), Foreign Policy (Mallery 1988), Problem Solving (Shapiro 1992), Judging a Turing Test (Dafna Shahaf and Amir March 26-28, 2007), Common Sense Knowledge (Andrich, Novosel et al. 2009), Speech Understanding (Dafna Shahaf and Amir March 26-28, 2007), Knowledge Representation and Reasoning (Shapiro 1992), Word Sense Disambiguation (Navigli and Velardi July 2005; Chen, et al. November 30, 2009), Machine Translation (Wikipedia Retrieved January 7, 2011), Ubiquitous Computing (Leahu, et al. September 21 - 24, 2008), Change Management for Biomedical Ontologies (Nejad April 2010), Natural Language Understanding (Shapiro 1992), Software Brittleness (Wikipedia Retrieved January 7, 2011), Vision or Image Understanding (Shapiro 1992).
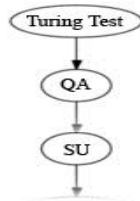

Figure 3.   Reductions from the first NP-Complete problem.

E. 1st AI-Hard Problem: Programming

We define the problem of Programming as taking a natural language description of a program and producing a source code which then compiled on some readily available hardware/software produces a computer program which satisfies all implicit and explicit requirements provided in the natural language description of the programming problem assignment. Simple examples of Programming are typical assignments given to students in computer science classes. Ex. "Write a program to play Tic-Tac-Toe." with successful students writing source code which if correctly compiled allows the grader to engage the computer in an instance of that game. Many requirements of such assignment remain implicit such as that response time of the computer should be less than a minute. Such implicit requirements are usually easily inferred by students who have access to culture instilled common sense. As of this writing no program is capable of solving Programming outside of strictly restricted domains.

Having access to an Oracle capable of solving Programming allows us to solve TT via a simple reduction. For any statement S presented during TT we can transform said statement into a programming assignment of the form: "Write a program which would respond to S with a statement indistinguishable from a statement provided by an average human" (A full transcript of the TT may also be provided for disambiguation purposes). Applied to the set of all possible TT statements this procedure clearly allows us to pass TT, however Programming itself is not in the set of AI-Problems as there are many instances of Programming which are not solvable by Human Oracles. For example "Write a program to pass Turing Test" is not known to be an AI-Problem under the proposed definition. Consequently, Programming is an AI-Hard problem.

III.   BEYOND AI-COMPLETENESS

The human oracle function presented in this paper assumes that the human being behind it has some assistance from the computer in order to process certain human unfriendly data formats. For example a binary string representing a video is completely impossible for a human being to interpret but could easily be played by a computer program in the intended format making it possible for a human to solve a video understanding related AI-Complete problem. It is obvious that a human being provided with access to a computer (perhaps with Internet connection) is more intelligent compared to a human unenhanced in such a way. Consequently it is important to limit help from a computer to a human worker inside a human Oracle function to assistance in the domain of input/output conversion but not beyond as the resulting function would be both AI-Complete and "Computer Complete".
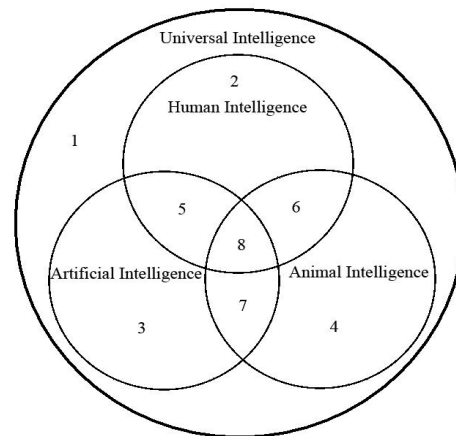

Figure 4.   **Fig. 1.** Venn diagram for four different types of intelligence.

Figure 4 utilizes a Venn diagram to illustrate subdivisions of problem space produced by different types of intelligent computational devices. Region 1 represents what is known as Universal Intelligence (Legg and Hutter December 2007) or Super Intelligence (R.V. Yampolskiy 2011; Roman V. Yampolskiy 2012; Roman V. Yampolskiy and Fox 2012b, 2012a; Legg June 2008; Roman V. Yampolskiy October 3-4, 2011a, October 3-4, 2011b) a computational agent which outperforms all other intelligent agents over all possible environments. Region 2 is the standard unenhanced Human level intelligence of the type capable of passing a Turing Test, but at the same time incapable of computation involving large numbers or a significant amount of memory. Region 3 is what is currently possible to accomplish via the state-of-the-art AI programs. Finally Region 4 represents an abstract view of animal intelligence. AI intelligence researchers strive to produce Universal Intelligence and it is certainly likely to happen given recent trends in both hardware and software developments and theoretical underpinning of the Church/Turing Thesis (AM Turing 1936). It is also likely, that if we are able to enhance human minds with additional memory and port them to a higher speed hardware we will essentially obtain a Universal Intelligence (Sandberg and Boström 2008).

While Universal Intelligence incorporates abilities of all the lower intelligences it is interesting to observe that Human, AI and Animal intelligences have many interesting regions of intersection. For example animal minds are as good as human minds at visual understanding of natural scenes. Regions 5, 6, and 7 illustrate common problem spaces between two different types of intelligent agents. Region 8 represents common problem solving abilities of humans, computers and animals. Understanding such regions of commonality may help us to better separate involved computational classes which are represented by abilities of a specific computational agent minus the commonalities with a computational agent with which we are trying to draw a distinction. For example CAPTCHA (Ahn, et al. 2003) type tests rely on the inability of computers to perform certain pattern recognition tasks with the same level of accuracy as humans to separate AI agents from Human agents. Alternatively a test could be devised to tell humans not armed with calculators from AIs by looking at the upper level of ability. Such a test should be easy to defeat once an effort is made to compile and formalize the limitations and biases of the human mind.

It is also interesting to consider the problem solving abilities of hybrid agents. We have already noted that a human being equipped with a computer is a lot more capable compared to an unaided person. Some recent research in Brain Computer Interfaces (Vidal 1973) provides a potential path for future developments in the area. Just as interestingly combining pattern recognition abilities of animals with symbol processing abilities of AI could produce a computational agent with a large domain of human like abilities (see work on RoboRats (Talwar, Xu et al. 2 May 2002) on monkey controlled robots (Nicolelis, Wessberg et al. 2000)). It is very likely that in the near future the different types of intelligent agents will combine to an even greater extent. While such work is under way we believe that it may be useful to introduce some additional terminology into the field of problem classification. For the complete space of problems we propose that the computational agents which are capable of solving a specific subset of such problems get to represent the set in question. Therefore we propose additional terms: "Computer-Complete" and "Animal-Complete" to represent computational classes solvable by such agents. It is understood that just like humans differ in their abilities so do animals and computers. Aggregation and averaging utilized in our human function could be similarly applied to the definition of the respective oracles. As research progresses common names may be needed for different combinations of regions from Figure 8 illustrating such concepts as Human-AI hybrid or Animal-Robot hybrid.

IV. CONCLUSIONS

Progress in the field of artificial intelligence requires access to well defined problems of measurable complexity. The theory of AI-Completeness aims to provide a base for such formalization. Showing certain problems to be AI-Complete/-Hard is useful for developing novel ways of telling computers from humans. Also, any problem shown to be AI-Complete would be a great alternative way of testing an artificial intelligent agent to see if it attained human level intelligence (Dafna Shahaf and Amir March 26-28, 2007).

REFERENCES

Ahn, Lv. (June 2006). Games With A Purpose. *IEEE Computer Magazine*, 96-98.

Ahn, Lv, Blum, M, Hopper, N, and Langford, J. (2003). *CAPTCHA: Using Hard AI Problems for Security.* Eurocrypt.

Andrich, C, Novosel, L, and Hrnkas, B. (2009). *Common Sense Knowledge*. Paper presented at the Information Search and Retrieval, Available at: http://www.iicm.tu-graz.ac.at/cguetl/courses/isr/uearchive/uews2009/Ue06-CommonSenseKnowledge.pdf.

Anusuya, MA, and Katti, SK. (2009). Speech Recognition by Machine: A Review. *International Journal of Computer Science and Information Security (IJCSIS), 6(3)*, 181-205.

Bajaj, V. (April 25, 2010). *Spammers Pay Others to Answer Security Tests*. Paper presented at the The New York Times.

Bergmair, R. (December 2004). *Natural Language Steganography and an ``AI-complete" Security Primitive*. Paper presented at the 21st Chaos Communication Congress, Berlin.

Bradford, PG, and Wollowski, M. (1995). A formalization of the Turing Test. *SIGART Bulletin, 6(4)*, 3-10.

Chan, T-Y. (2003). *Using a Text-to-Speech Synthesizer to Generate a Reverse Turing Test.* 15th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'03).

Chen, J, Liu, J, Yu, W, and Wu, P. (November 30, 2009). *Combining Lexical Stability and Improved Lexical Chain for Unsupervised Word Sense Disambiguation*. Paper presented at the Second International Symposium on Knowledge Acquisition and Modeling (KAM '09), Wuhan

Demasi, P, Szwarcfiter, JL, and Cruz, AJO. (March 5-8, 2010). *A Theoretical Framework to Formalize AGI-Hard Problems*. Paper presented at The Third Conference on Artificial General Intelligence, Lugano, Switzerland.

Dimmock, N, and Maddison, I. (December 2004). Peer-to-peer collaborative spam detection. *Crossroads, 11(2)*.

Gentry, C, Ramzan, Z, and Stubblebine, S. (2005). *Secure distributed human computation*. Paper presented at the 6th ACM conference on Electronic commerce.

Hendler, J. (September 2008). We've Come a Long Way, Maybe …. *IEEE Intelligent Systems, 23(5)*, 2-3.

Hirschman, L, and Gaizauskas, R. (2001). Natural Language Question Answering. The View from Here. *Natural Language Engineering, 7(4)*, 275-300.

Horvitz, E. (2007). Reflections on Challenges and Promises of Mixed-Initiative Interaction. *AI Magazine-Special Issue on Mixed-Initiative Assistants, 28(2)*.

Horvitz, E, and Paek, T. (2007). Complementary Computing: Policies for Transferring Callers from Dialog Systems to Human Receptionists. *User Modeling and User Adapted Interaction, 17(1)*, 159-182.

Ide, N, and Véronis, J. (1998). Introduction to the special issue on word sense disambiguation: the state of the art. *Computational Linguistics, 24(1)*, 1-40.

Kapoor, A, Tan, D, Shenoy, P, and Horvitz, E. (2008). *Complementary Computing for Visual Tasks: Meshing Computer Vision with Human Visual Processing*. Paper presented at the IEEE International Conference on Automatic Face and Gesture Recognition.

Karp, RM. (1972). Reducibility Among Combinatorial Problems. In RE Miller & JW Thatcher (Eds.), *Complexity of Computer Computations* (pp. 85-103). New York: Plenum.

Leahu, L, Sengers, P, and Mateas, M. (September 21 - 24, 2008). *Interactionist AI and the promise of ubicomp, or, how to put your box in the world without putting the world in your box*. Paper presented at the Tenth International Conference on Ubiquitous Computing, Seoul, South Korea.

Legg, S. (June 2008). *Machine Super Intelligence*. Paper presented at the PhD Thesis, University of Lugano, Available at: http://www.vetta.org/documents/Machine_Super_Intelligence.pdf.

Legg, S, and Hutter, M. (December 2007). Universal Intelligence: A Definition of Machine Intelligence. *Minds and Machines, 17(4)*, 391-444.

Mallery, JC. (1988). *Thinking About Foreign Policy: Finding an Appropriate Role for Artificially Intelligent Computers*. Paper presented at the Annual Meeting of the International Studies Association, St. Louis, MO.

McDaniel, R, and Yampolskiy, RV. (2011). *Embedded non-interactive CAPTCHA for Fischer Random Chess*. Paper presented at the 16th International Conference on Computer Games (CGAMES), Louisville, KY.

McIntire, JP, Havig, PR, and McIntire, LK. (July 21-23, 2009). *Ideas on authenticating humanness in collaborative systems using AI-hard problems in perception and cognition*. Paper presented at the IEEE National Aerospace & Electronics Conference (NAECON), Dayton, OH.

McIntire, JP, McIntire, LK, and Havig, PR. (May 18-22, 2009). *A variety of automated turing tests for network security: Using AI-hard problems in perception and cognition to ensure secure collaborations*. Paper presented at the International Symposium on Collaborative Technologies and Systems (CTS '09) Baltimore, MD.

Mert, E, and Dalkilic, C. (September 14-16, 2009). *Word sense disambiguation for Turkish*. Paper presented at the 24th International Symposium on Computer and Information Sciences (ISCIS 2009), Guzelyurt.

Morgan, N, Baron, D, Bhagat, S, Carvey, H, Dhillon, R, Edwards, J, . . . Wooters, C. (April 6-10, 2003). *Meetings about meetings: research at ICSI on speech in multiparty conversations*. Paper presented at the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03).

Mueller, ET. (March 1987). *Daydreaming and Computation. Ph.D. Dissertation, University of California*. Los Angeles.

Navigli, R, and Velardi, P. (July 2005). Structural Semantic Interconnections: A Knowledge-Based Approach to Word Sense Disambiguation. *IEEE Transactions On Pattern Analysis and Machine Intelligence, 27(7)*, 1075-1086.

Nejad, AS. (April 2010). *A Framework for Analyzing Changes in Health Care Lexicons and Nomenclatures. PhD dissertation. Concordia University*. Montreal, Quebec, Canada.

Nicolelis, MAL, Wessberg, J, Stambaugh, CR, Kralik, JD, Beck, PD, Laubach, M, . . . Kim, J. (2000). Real-

time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature, 408(6810)*, 361.

Pepitone, J. (Retrieved on: January 13, 2011). *IBM's Jeopardy supercomputer beats humans in practice bout*. Paper presented at the CNNMoney, Available at: http://money.cnn.com/2011/01/13/technology/ibm_jeopardy_watson.

Phillips, PJ, and Beveridge, JR. (September 28-30,. 2009). *An introduction to biometric-completeness: The equivalence of matching and quality*. Paper presented at the IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems (BTAS '09) Washington, DC

Raymond, ES. (March 22, 1991). *Jargon File Version 2.8.1*. Available at: http://catb.org/esr/jargon/oldversions/jarg282.txt.

Salloum, W. (November 30, 2009). *A Question Answering System based on Conceptual Graph Formalism*. Paper presented at the The 2nd International Symposium on Knowledge Acquisition and Modeling (KAM 2009), China.

Sandberg, A, and Boström, N. (2008). *Whole Brain Emulation: A Roadmap*. Paper presented at the Future of Humanity Institute, Oxford University. Technical Report #2008-3, Available at: http://www.fhi.ox.ac.uk/Reports/2008-3.pdf.

Searle, J. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences, 3(3)*, 417-457.

Shahaf, D, and Amir, E. (March 26-28, 2007). *Towards a theory of AI completeness*. Paper presented at the 8th International Symposium on Logical Formalizations of Commonsense Reasoning (Commonsense 2007), California.

Shahaf, D, and Horvitz, E. (July 2010). *Generalized Task Markets for Human and Machine Computation*. Paper presented at the Twenty-Fourth AAAI Conference on Artificial Intelligence, Atlanta, GA.

Shapiro, SC. (1992). Artificial Intelligence. In SC Shapiro (Ed.), *Encyclopedia of Artificial Intelligence* (pp. 54-57). New York: John Wiley.

Shieber, SM. (December 2007). The Turing Test as Interactive Proof. *Nous, 41(4)*, 686-713.

Shieber, SM. (July 16-20, 2006). *Does the Turing Test demonstrate intelligence or not*. Paper presented at the Twenty-First National Conference on Artificial Intelligence (AAAI-06), Boston, MA.

Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*: Little, Brown.

Takagi, H. (2001). Interactive Evolutionary Computation: Fusion of the Capacities of EC Optimization and Human Evaluation. *Proceesings of the IEEE 89, 9*, 1275-1296.

Talwar, SK, Xu, S, Hawley, ES, Weiss, SA, Moxon, KA, and Chapin, JK. (2 May 2002). Behavioural neuroscience: Rat navigation guided by remote control. *Nature, 417*, 37-38.

Taylor, P, and Black, A. (1999). *Speech synthesis by phonological structure matching*. In *Eurospeech99*, Budapest, Hungary.

Turing, A. (1950). Computing Machinery and Intelligence. *Mind, 59(236)*, 433-460.

Turing, AM. (1936). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society, 42*, 230-265.

Vaas, L. (December 1, 2007). *Striptease Used to Recruit Help in Cracking Sites*. Paper presented at the PC Magazine.

Vidal, J. (1973). Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering, 2*, 157-180.

Wikipedia. (Retrieved January 7, 2011). *AI-Complete*. Available at: http://en.wikipedia.org/wiki/AI-complete.

Yampolskiy, RV. (2007a, April 13, 2007). *Embedded CAPTCHA for Online Poker*. 20th Annual CSE Graduate Conference (Grad-Conf2007), Buffalo, NY.

Yampolskiy, RV. (2007b, September 28, 2007). *Graphical CAPTCHA embedded in cards.* Western New York Image Processing Workshop (WNYIPW) - IEEE Signal Processing Society, Rochester, NY.

Yampolskiy, RV. (2011). AI-Complete CAPTCHAs as Zero Knowledge Proofs of Access to an Artificially Intelligent System. *ISRN Artificial Intelligence, 271878*.

Yampolskiy, RV. (2012). Leakproofing Singularity - Artificial Intelligence Confinement Problem. *Journal of Consciousness Studies (JCS), 19(1-2)*, 194–214.

Yampolskiy, RV. (October 3-4, 2011a). *Artificial Intelligence Safety Engineering: Why Machine Ethics is a Wrong Approach*. Paper presented at the Philosophy and Theory of Artificial Intelligence (PT-AI2011), Thessaloniki, Greece.

Yampolskiy, RV. (October 3-4, 2011b). *What to Do with the Singularity Paradox?* Paper presented at the Philosophy and Theory of Artificial Intelligence (PT-AI2011), Thessaloniki, Greece.

Yampolskiy, RV, and Fox, J. (2012a). Artificial Intelligence and the Human Mental Model. In A Eden, J Moor, J Soraker & E Steinhart (Eds.), *In the Singularity Hypothesis: a Scientific and Philosophical Assessment*: Springer.

Yampolskiy, RV, and Fox, J. (2012b). Safety Engineering for Artificial General Intelligence. *Topoi Special Issue on Machine Ethics & the Ethics of Building Intelligent Machines, (In Press)*.

Yampolskiy, RV, and Govindaraju, V. (2007). Embedded Non-Interactive Continuous Bot Detection. *ACM Computers in Entertainment, 5(4)*, 1-11.