What Do You Mean by "AI"?

Pei Wang
Temple University, Philadelphia, USA
http://www.cis.temple.edu/~pwang/

Abstract

Many problems in AI study can be traced back to the confusion of different research goals. In this paper, five typical ways to define AI are clarified, analyzed, and compared. It is argued that though they are all legitimate research goals, they lead the research to very different directions, and most of them have trouble to give AI a proper identity. Finally, a solution is suggested.

1 The Problem of AI

"AI has always been a strange field" [Allen 1998], where researchers have very different opinions on fundamental issues [Kirsh 1991, Hearst and Hirsh 2000]. People not only argue on what is the best solution to the problem (which is normal), but also on what the problem is (which is unusual, at least when compared with the fields in computer science). At the 50th anniversary of the field, the AAAI Presidential Address still asked the question "what really is AI and what is intelligence about?" [Brachman 2006].

It is well known that the phrase "Artificial Intelligence (AI)" means different things to different people. However, this issue has not been explored to the extent it deserves. This paper will clarify various perspectives of AI, analyze their relations, and evaluate their potentials, especially the identities they give to AI, respectively.

Though people have different opinions on how to accurately define "Artificial Intelligence", on a more general level they do agree on what this field is about. Human beings differ from animals and machines mainly in their mental, or cognitive, ability, which is commonly called "intelligence", and AI is the attempt to reproduce this ability in computer systems. This vague consensus sets important constraints on how AI should be defined:

- Since the best example of "intelligence" comes from the human mind, AI should be defined as similar to human intelligence in certain aspects.
- Since AI is an attempt to duplicate human intelligence, not to completely duplicate a human being, AI should be defined as different from human

intelligence in certain aspects. Otherwise it would be about "artificial person", rather than intelligent computer.

- AI should not be defined in such a narrow way that takes human intelligence as the only possible form of intelligence, otherwise AI research would be impossible.
- AI should not be defined in such a broad way that takes existing computer systems as already having intelligence, otherwise AI research would be unnecessary.

Under the above constraints, the major difference among various AI definitions is in the aspects of human intelligence that must be duplicated, and implicitly, in the aspects of human intelligence that can be omitted.

To make the analysis and comparison precise, in this paper both human beings and computer systems are described as "agents" that "receive percepts from the environment and perform actions" [Russell and Norvig 2002]. At a given moment t, the full history of an agent can be represented as $<P, S, A>$, where $P = <p_0, \cdots, p_t>$ is the stream of percepts, $A = <a_0, \cdots, a_t>$ is the stream of actions, and $S = <s_0, \cdots, s_t>$ is the sequence of internal states the system has gone through. When a typical human mind is represented as $H = <P^H, S^H, A^H>$, and a typical intelligent computer as $C = <P^C, S^C, A^C>$, we will see that different ways to define AI correspond to different senses in which C is *similar* to H.

2 Typical Ways to Define AI

According to the distinction made in [Wang 1994], there are five typical ways to define AI, corresponding to establish the similarity between C and H by *structure*, *behavior*, *capability*, *function*, and *principle*, respectively.

By Structure

Since the best known instance of intelligence is produced by the human brain, it is natural to assume that AI can be achieved by building a brain-like *structure*, consisting of neuron-like processing units working in parallel.

This idea has been implemented in various forms, such as Connection Machine [Hillis 1986] and Artificial Neural Networks [Smolensky 1988]. More recent brain-oriented works include [Hawkins and Blakeslee 2004, Markram 2006].

Due to the complexity of the human brain and its fundamental difference from computer hardware, none of these projects plan to be faithful to the brain structure in all the details. Instead, they only take the brain as the source of inspirations, and the resulting systems only approximate to the brain at a certain level and scope of description.

Even so, many people inside and outside the field of AI believe that accurate "brain modeling" will provide the ultimate solution to AI, when it is allowed by our knowledge of the human brain and the available computer technology. According to this opinion, "the ultimate goals of AI and neuroscience are quite similar" [Reeke and Edelman 1988].

Let us call this type of definition "**Structure-AI**", since it requires the structural similarity between an AI system and the human brain. In the agent framework, it means that C is similar to H in the sense that

$$<P^C, S^C, A^C> \approx <P^H, S^H, A^H>$$

that is, the two should have similar streams of percepts and actions, as well as similar state transforming sequences, due to their similar internal structure.

According to this understanding of AI, even though it is impossible to accurately duplicate the brain structure in the near future, we should try to move to that goal as close as possible, and the distance to it can be used to evaluate the research results.

By Behavior

Since intelligence seems to be more about the human mind than the human brain, many people believe that it is better to concentrate on the system's *behavior* when evaluating its intelligence. The best known idea in this category is the Turing Test [Turing 1950].

Though Turing proposed his test only as a sufficient condition, not a necessary condition, for intelligence, it nevertheless has been taken by many people as the definition of AI [Brackenbury and Ravin 2002, Schubert 2006].

A typical opinion can be found in Newell's discussion of the Soar project [Newell 1990], which is presented both as an AI system and a model of human psychology. According to this opinion, AI is identified with "cognitive modeling", where the results are evaluated by comparisons with psychological data produced by human subjects.

Another example of this understanding of AI is in the field of "chatbot", where the intelligence of a system is evaluated according to how much it "talks like a human", such as in the Loebner Prize Competition [Mauldin 1994].

Let us call this type of definition "**Behavior-AI**", since it requires the behavioral similarity between an AI system and the human mind. In the agent framework, it means that C is similar to H in the sense that

$$<P^C, A^C> \approx <P^H, A^H>$$

that is, the two should have similar streams of percepts and actions. Here the two systems are taken to be "black box", whose internal structure and state do not matter. Of course, the AI system may be similar to a human mind only after a certain period of training, and that can be accepted in the above representation by setting the starting moment of the percepts and actions after the completion of the training.

By Capability

For people whose interest in AI mainly comes from its potential practical applications, the intelligence of a system should be indicated by its *capability* of solving hard problems [Minsky 1985]. After all, this is how we judge how intelligent a person is. Also, the progress of a research field will eventually be evaluated according to the usefulness of its results.

Partly because of such considerations, the earliest practical problems studied by AI were typical intellectual activities like theorem proving and game playing — if a person can solve these problems, we call the person "intelligent"; therefore, if a computer can do the same, then we may have to call the computer "intelligent", too. Driven by similar motivations, a large number of application-oriented AI projects are "expert systems" in various domain — experts are intelligent, so if a computer can solve a problem that only an expert can, the computer must be intelligent, too.

Especially, a computer is often considered as intelligent if it solves a problem that could only be solved by human beings previously. Consequently, AI becomes an expanding frontier of computer application.

The biggest AI achievements, according to this understanding, include Deep Blue, the computer that defeated the world champion in chess, and Stanley, the self-driven vehicle that finished a 132-mile trek in 7 hours. Here is a recent form of this idea: "I suggest we replace the Turing test by something I will call the 'employment test'. To pass the employment test, AI programs must be able to perform the jobs ordinarily performed by humans. Progress toward human-level AI could then be measured by the fraction of these jobs that can be acceptably performed by machines." [Nilsson 2005]

Let us call this type of definition "**Capability-AI**", since it requires an AI system to have human capability of practical problem solving. In the agent framework, it means that C is similar to H in the sense that there are moments i and j that

$$<p_i^C, a_i^C> \approx <p_j^H, a_j^H>$$

that is, the action (solution) the computer produces for a percept (problem) is similar to the action produced by a human to a similar percept. To make discussion simple, here we assume that a single percept can represent the problem, and a single action can represent the solution.

Since here what matters is the final solution only, it is irrelevant whether the computer goes through a human-like internal process or produce human-like external behavior beyond this problem-solving process. It follows that systems with higher intelligence can solve more and harder problems, as suggested by Nilsson.

By Function

Since most AI researchers are computer scientists and engineers, they prefer to represent the ability of an agent as some *function* that maps input (percepts) into output (actions), which is how a computer program is specified.

Typical opinions are like "Intelligence is the computational part of the ability to achieve goals in the world" and "What is important for AI is to have algorithms as capable as people at solving problems", both from [McCarthy 2004]. A more systematic and influential description came from Marr: "a result in Artificial Intelligence consists of the isolation of a particular information processing problem, the formulation of a computational theory for it, the construction of an algorithm that implements it, and a practical demonstration that the algorithm is successful." [Marr 1977]

Guided by such opinions, the field of AI is widely seen as consisting of separate cognitive functions, such as searching, reasoning, planning, learning, communicating, perceiving, acting, etc., each having its various forms of computational definitions and algorithmic implementations. [Russell and Norvig 2002]

Let us call this type of definition "**Function-AI**", since it requires an AI system to have human cognitive functions. In the agent framework, it means that C is similar to H in the sense that there are moments i and j that

$$a_i^C = f^C(p_i^C)\,,\ a_j^H = f^H(p_j^H)\,,\ \text{ and }\ f^C \approx f^H$$

that is, the function that maps a percept (problem) into an action (solution) in the computer is similar to that of a human. Since here the focus is on the functions, the actual percepts and actions of the two agents do not have to be similar to each other. Naturally, a system with higher intelligence should implement more such functions efficiently.

By Principle

Science always looks for simple and unified explanations of complicated and diverse phenomena. Therefore, it is not a surprise that some AI researchers attempt to identify the fundamental *principle* by which human intelligence can be explained and reproduced in computer at a general level.

Intuitively, "intelligence" is associated with the ability to get the best solution. However, such a definition would be trivial if the agent could exhaustively evaluate all possible solutions and select the best among them. To be more realistic, Simon proposed the notion of "Bounded Rationality", which restricts what the agent can know and do [Simon 1957]. Russell argued that intelligent agents should have "Bounded Optimality", the ability to generate maximally successful behavior given the available information and computational resources [Russell 1997].

As a concrete example of such an idea, the NARS project has been carried out according to the belief that "intelligence" means "adaptation with insufficient knowledge and resources", which requires the system to be finite, real-time, and open [Wang 2006]. Consequently, NARS is designed under stronger restriction than imposed by "Bounded Rationality" and "Bounded Optimality".

Let us call this type of definition "**Principle-AI**", since it requires an AI system to follow similar principle as the human mind. In the agent framework, it means that C is similar to H in the sense that

$$A^C = F^C(P^C)\,,\ A^H = F^H(P^H)\,,\ \text{ and }\ F^C \approx F^H$$

that is, the function that maps the whole stream of percepts into the whole stream of actions in the computer is similar to that of a human. Again, here the focus is on the function, not the actual percepts and actions. The function is called a "principle", to stress that it is not merely about a single problem and its solution, but about the agent's life-long history in various situations, when dealing with various types of problems.

3 The Necessity of Distinction

The above five definitions of AI are all legitimate research goals, which are different from each other.

Structure-AI mainly contributes to the study of the human brain. It also helps to explain how the brain carries out various cognitive activities, but if the goal is in the behavior, capability, function, or principle of the mind, then to duplicate the brain structure is often not the best way (in terms of simplicity and efficiency), because the brain is formed under biological and evolutionary restrictions, which are largely irrelevant to computers.

Behavior-AI mainly contributes to the study of the human psychology. Very often, "the human way" give us inspirations on how to use a computer, but it is not the best way to solve a practical problem, or to implement a cognitive function or principle. Also, behavior similarity does not mean structural similarity.

Capability-AI mainly contributes to various application domains, by solving practical problems there. However, due to the lack of generality of the solutions, this kind of solution usually contributes little to the study of brain or mind outside the domain.

Function-AI mainly contributes to computer science, by producing new software (sometimes also hardware) that can carry out a certain type of computation. However, the best way to implement the required computation is usually not exactly the way such a process is carried out in the human mind/brain complex. Since a cognitive function is generalized over many concrete problems, it is not necessarily the best way to solve each of them. If an agent is equipped with multiple cognitive functions, they are not necessarily designed according to the same principle.

Principle-AI mainly contributes to the study of information processing in different situations, by exploring the implications of different assumptions. Given the generality of a principle, it cannot explain all the details of the human brain or the human mind, nor does it provide the best way to solve every practical problem. Even though a principle-based system usually does carry out various cognitive functions, they are not necessarily separate processes, each with its computational definition and algorithmic implementation.

In summary, they are not five trails to the same summit, but to five different summits.

In [Russell and Norvig 2002], the definitions of AI are organized into four categories, according to two distinctions: (1) "humanly" vs. "rationally", and (2) "think" vs. "act". How are these four categories related to the above five?

According to the above analysis, their first distinction is not very precise — though "Structure-AI" and "Behavior-AI" are clearly on the "humanly" side, the other three are also "humanly" (though at different levels or scopes), as well as "rationally" (again, in different senses).

Their second distinction is hard to justify, because every agent has an internal mechanism that can be roughly referred to as "thinking" and an input/output mechanism that can be roughly referred to as "acting", and they must work consistently, no matter which of the two is mentioned in a definition. For the difference between "act humanly" and "think humanly", their examples are "Turing Test" and "Cognitive Modeling", respectively. However, both Turing Test and Cognitive Modeling require "intelligence" to be evaluated by behavior, though in Cognitive Modeling the internal process (thinking) is explicitly specified. For the difference between "think rationally" and "act rationally", their examples are "Logicist AI" and "Rational Agent", respectively. However, we cannot say that a logicist system does not act, or that an agent does not think. Their discussion about the difference between the two are actually about different types of rationalities.

In summary, there are reasons to believe that the five categories introduced above is a better partition than the four provided in [Russell and Norvig 2002].

To distinguish five types of AI definitions does not mean that they are not related to each other. It is possible to accept a definition of AI as the primary goal, and also to achieve some secondary goals at the same time, or to benefit from works aimed at a different goal. For example, when implementing a principle, we may find that the "human way" is very simple and efficient, which also provides good solutions to some real-world problems. However, even in such a situation, it is still necessary to distinguish the primary goal of a research from the additional and secondary results it may produce, because whenever there is a conflict (which is the usual case, rather than exceptional), it is the primary goal that should be used to justify the design decision.

Even though each of the five types of AI definition is valid, to mix them together in one project is not a good idea. Many current AI projects have no clearly specified goal, and people working on them often swing among different AI definitions. Such a practice causes inconsistency in the criteria of design and evaluation.

It is true that in many science disciplines the basic notions become well-defined only after long-term research, but in those disciplines, at least the phenomena to be studied are clearly identified at the beginning. On the contrary, in AI each researcher has to decide, at the very beginning, *which aspects* of the human intelligence should be studied, which is based on a explicitly or implicitly accepted working definition of AI. There is no way to be "definition-neutral", because otherwise the research would have nowhere to start — a phenomenon is relevant to AI only when the term "AI" has meaning, no matter more vague or poor the meaning is.

The confusion among different AI definitions is also a common root of many controversies in AI. For example, there has been a debate on whether Deep Blue is a success of AI [Allen 1998, McDermott 2001]. According to the above analysis, the conclusion should clearly be "yes" if "AI" is taken to mean "Capability-AI", otherwise the answer should be "not much", or even "no".

A common mistake is to believe that there is a "true" ("real", "natural") meaning of "intelligence" that AI must follow, though it comes in different forms:

Some people think that AI should follow the common usage (i.e., the dictionary definition) of the word "intelligence". This is not going to work. The meaning of "intelligence" in English (or a similar word in another natural language) was mostly formed before AI time, and therefore is mainly about human intelligence, where the various aspects (structure, behavior, capability, function, principle, etc.) are unified. On the contrary, for computer systems these aspects become different goals, as discussed previously.

For similar reasons, AI cannot simply borrow the definition of "intelligence" from other disciplines, such as psychology or education, though the notion does have a longer history in those fields. This is not only because there are also controversies in those domains about what intelligence is, but also because the notion "intelligence" is mainly used there to stress *the difference among human beings* in cognitive ability. On the contrary, for AI this difference is almost neglectable, and the notion is mainly used to stress *the difference between human beings and computer systems*. For this reason, it is not a good idea to use IQ test to judge the ability of AI systems.

Some people argue that "AI is what the AI researchers do". Though a survey of the field provides a valid *descriptive definition* of AI, it is not a valid *working definition*, which should be precise and coherent to guide a research project [Wang 1994]. Under the common name "AI", AI researcher are actually doing quite different things, as shown previously. Even if there is a majority point of view, it does not necessary become the "true meaning" of AI that everyone must follow.

4 The Possibility of Comparison

To say that all the five types of working definitions are valid does not mean that they cannot be compared with respect to certain criteria.

In [Wang 1994], four criteria of a good working definition were borrowed from Carnap's work on the notion of "probability" [Carnap 1950]:

- It should have a *sharp* boundary.

- It should be *faithful* to the notion to be clarified.
- It should lead to *fruitful* research.
- It should be as *simple* as possible.

Given their forms as defined previously, the five types of definition are similar with respect to the requirements of sharpness and simplicity. Therefore, the following discussion will be focused on the other two criteria.

As analyzed before, in general it is hard to say which of the five is more faithful to the everyday usage of the word "intelligence", because each of them capture a different aspect of it. However, for AI to be established as a discipline of science, the working definition needs to satisfy some special constraints. As argued at the beginning of the paper, the extension of "intelligent system" should neither be so narrow as to include human beings only, nor be so broad as to include all existing computer systems.

Similarly, though each of the five ways to define AI leads to fruitful research, we can compare them with respect to their ability to give AI a proper *identity*, which should explain how the field differs from the other disciplines, as well as elicits the common natures of all AI projects.

The above requirements are not arbitrarily selected. AI has been suffering from a serious identity crisis for years. Many AI researchers have complained that the field has not got the credit it deserves, which is called "The AI Effect" — as soon as a problem is solved, it is no longer considered as "AI" anymore [Schank 1991]. Within the field, fragmentation is also a big problem [Brachman 2006] — each subfield has its own goal and methods, and to collectively call them "AI" seems only have historical reason, that is, they all more or less come out of attempts of making computer "intelligent", whatever it means. To label them with a umbrella term "intelligence" does little to improve the situation, since the term lacks substance to bound the subfields together.

Now let us analyze the responsibility of each type of working definition with respect to the identity problem the field AI faces. Especially, how *AI* is related to *human intelligence* and *computer science*, respectively.

There is no doubt that the best example of "intelligence" is "human intelligence", and therefore all working definitions attempt to make computer systems similar to humans, in various senses, and to various extents. However, **Structure-AI** and **Behavior-AI** seem to leave too little space for "non-human intelligence" — they may be sufficient conditions for "intelligence", but unlikely to be necessary conditions. If an intelligent system must have human brain structure or produce human cognitive behaviors, then some other possibilities, such as "animal intelligence", "collective (group) intelligence", and "extraterrestrial intelligence" all become impossible *by definition.* It would be similar to defining "vision" by the structure or behavior of human visual organ. For AI, such a definition will seriously limit our imagination and innovation of novel forms of intelligence. Human intelligence is developed under certain evolutionary and biological restrictions, which are essential for human, but hardly for intelligence in general. After all, "Artificial Intelligence" should not be taken to mean "Artificial Human Intelligence", since "Intelligence" should be more general than "Human Intelligence".

On the other hand, **Capability-AI** and **Function-AI** seem to allow too many systems to be called "intelligent". It is not hard to recognize that works under the former is just like what we usually call "computer application", and the latter, "computer science", except that the problems or tasks are those that "humans can do or try to do" [Allen 1998]. Do these definitions give enough reason to distinguish AI from Computer Science (CS)? Marr's computation-algorithm-implementation analysis of AI [Marr 1977] can be applied to every problem studied in CS, and so does the following textbook definition: "we define AI as the study of agents that receive percepts from the environment and perform actions" [Russell and Norvig 2002]. This consequence is made explicit by the claim of Hayes and Ford that AI and CS are the same thing [Hayes and Ford 1995].

If the only difference between AI and CS is that the "AI problems" are historically solved by the human mind, then how about problems like sorting or evaluating arithmetic expression? Some people have argued that every program is intelligent, and "intelligence" is a matter of degree. Such a usage of the concept of "intelligence" is coherent, except that the concept has been trivialized too much. If this is the case, there is no wonder why AI has got little credit and recognition — if everything developed in the field of AI can be done in CS, and "intelligent agent" has no more content than "agent", what difference does it make if we omit the fancy label "intelligence"?

Furthermore, the widely acceptance of **Capability-AI** and **Function-AI** are responsible for fragmentation of AI. Roughly speaking, the former is behind most of the "applied AI" works, and the latter, "theoretical AI" works. Both of them defining AI by a *group* (of capabilities and functions, respectively), without demanding much commonality among its members. As a result, AI practitioners usually assume they can, and should, start to work on a single capability or function, which may be integrated to get a general intelligence in the future. Since the best ways to solve a practical problem or to carry out a formal computation differ greatly from case to case, there is not too much to be learned from each other, even though all of them are called "AI". As far as people continue to define their problems in this way, the fragmentation will continue, too.

The above analysis leaves us only with **Principle-AI**. Of course, like the other four types discussed above, Principle-AI is not a single working definition of AI, but a group of them. Different members in the group surely lead to different consequences. Obviously, if the "principle" under consideration is too broad, it will include all computer systems (which is bad); if it is too narrow, it will exclude all non-human systems (which is bad, too). Therefore we need something *in between*, that is, a principle that (1) is followed by the human mind, (2) can be followed by computer systems, (3) but are not followed by traditional computer systems.

An example of such a working definition of AI is the one accepted in the NARS project. Briefly speaking, it identifies "intelligence" with "adaptation with insufficient knowledge and resources", which implies that the system is finite, works in real-time, is open to novel tasks, and learns from experience [Wang 1994, Wang 2006]. There are many reasons to believe that the human mind is such a system. The practice of NARS shows that it is possible to develop a computer system following this principle. Finally, traditional computer systems do not follow this principle. Therefore, such a working definition satisfies the previous requirements. Furthermore, though NARS can be studied in different aspects, the system cannot be divided into independent functions or capabilities, since all components of the system are designed according to the same principle [Wang 2006]. The notion of "intelligence" is not an optional label in this research, since it does introduce ideas not available in computer science or cognitive psychology.

To prefer the NARS definition of AI does not mean that it can replace the others for all purposes. As discussed before, each valid working definition of AI has its value. Principle-based definitions are often described as "looking for a silver bullet", labeled as "physics envy", and rejected by arguments like "intelligence is too complicated to be explained by a few simple principles". However, all these criticisms take such a definition (of Principle-AI) as the *means* to achieve other *ends* (Structure-AI, Behavior-AI, Capability-AI, or Function-AI), which is a misconception. The NARS definition may give AI a better identity than the other definitions do, though the former does not produce the values that can be produced by the latter.

Obviously, the NARS definition of AI is not a *descriptive definition* of the term, that is, its common use in the field. On the contrary, most of the existing "AI systems" do not satisfy this definition. However, it does not mean that the definition should be rejected, but that the field should change into a more coherent and fruitful discipline of science.

5 Conclusion

Though intuitively everyone agree that AI means to build computer systems that are similar to the human mind, they have very different ideas on where this similarity should be. Typical opinions define this similarity in terms of *structure*, *behavior*, *capability*, *function*, and *principle*.

These working definitions of AI are all valid, in the sense that each of them captures an aspect of human intelligence, and makes it as a precise research goal, which is achievable to various extent, as far as the "similar to" human intelligence in the definition is not replaced by "identical to". Each of them are also fruitful, in the sense that it has guided the research to produce results with intellectual or practical values.

On the other hand, these working definitions are different, in the sense that they set different goals, require different methods, produce different results, and evaluate progress according to different criteria. They cannot replace one another.

The current AI research suffers from the confusion of various goals and the missing of an identity. Consequently, many debates are caused by misunderstanding, and the field as a whole is fragmented within, as well as has trouble to justify its uniqueness and integrity to the outside world.

To solve these problems, the most promising way is to define AI by a principle of rationality that is followed by the human mind, but not by traditional computer systems. The NARS project shows that this is a possible solution.

References

[Allen 1998] Allen, J. F. (1998). AI growing up: The changes and opportunities. *AI Magazine*, 19(4):13–23.

[Brachman 2006] Brachman, R. J. (2006). (AA)AI — more than the sum of its parts. *AI Magazine*, 27(4):19–34.

[Brackenbury and Ravin 2002] Brackenbury, I. F. and Ravin, Y. (2002). Machine intelligence and the Turing Test. *IBM Systems Journal*, 41(3):524–529.

[Carnap 1950] Carnap, R. (1950). *Logical Foundations of Probability.* The University of Chicago Press, Chicago.

[Hawkins and Blakeslee 2004] Hawkins, J. and Blakeslee, S. (2004). *On Intelligence*. Times Books, New York.

[Hayes and Ford 1995] Hayes, P. and Ford, K. (1995). Turing test considered harmful. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pages 972–977.

[Hearst and Hirsh 2000] Hearst, M. A. and Hirsh, H. (2000). AI's greatest trends and controversies. *IEEE Intelligent Systems*, pages 8–17.

[Hillis 1986] Hillis, W. D. (1986). *The Connection Machine*. MIT Press, Cambridge, Massachusetts.

[Kirsh 1991] Kirsh, D. (1991). Foundations of AI: the big issues. *Artificial Intelligence*, 47:3–30.

[Markram 2006] Markram, H. (2006). The Blue Brain project. *Nature Reviews Neuroscience*, 7(2):153–160.

[Marr 1977] Marr, D. (1977). Artificial intelligence: a personal view. *Artificial Intelligence*, 9:37–48.

[Mauldin 1994] Mauldin, M. L. (1994). ChatterBots, TinyMuds, and the Turing Test: entering the Loebner Prize competition. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 16–21.

[McCarthy 2004] McCarthy, J. (2004). What is Artificial Intelligence? URL: *www-formal.stanford.edu/jmc/whatisai/whatisai.html*.

[McDermott 2001] McDermott, D. (2001). *Mind and Mechanism.* MIT Press, Cambridge, Massachusetts.

[Minsky 1985] Minsky, M. (1985). *The Society of Mind.* Simon and Schuster, New York.

[Newell 1990] Newell, A. (1990). *Unified Theories of Cognition.* Harvard University Press, Cambridge, Massachusetts.

[Nilsson 2005] Nilsson, N. J. (2005). Human-level artificial intelligence? Be serious! *AI Magazine*, 26(4):68–75.

[Reeke and Edelman 1988] Reeke, G. N. and Edelman, G. M. (1988). Real brains and artificial intelligence. *Dædalus*, 117(1):143–173.

[Russell 1997] Russell, S. (1997). Rationality and intelligence. *Artificial Intelligence*, 94:57–77.

[Russell and Norvig 2002] Russell, S. and Norvig, P. (2002). *Artificial Intelligence: A Modern Approach.* Prentice Hall, Upper Saddle River, New Jersey, 2nd edition.

[Schank 1991] Schank, R. C. (1991). Where is the AI? *AI Magazine*, 12(4):38–49.

[Schubert 2006] Schubert, L. K. (2006). Turing's dream and the knowledge challenge. In *Proceedings of the Twenty-first National Conference on Artificial Intelligence*, pages 1534–1538, Menlo Park, California. AAAI Press.

[Simon 1957] Simon, H. A. (1957). *Models of Man: Social and Rational.* John Wiley, New York.

[Smolensky 1988] Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 11:1–74.

[Turing 1950] Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, LIX:433–460.

[Wang 1994] Wang, P. (1994). On the working definition of intelligence. Technical Report 94, Center for Research on Concepts and Cognition, Indiana University, Bloomington, Indiana.

[Wang 2006] Wang, P. (2006). *Rigid Flexibility: The Logic of Intelligence.* Springer, Dordrecht.