John Haugeland (1981) – Semantic Engines: An Introduction to Mind Design
Why is the idea that minds are digital computers so appealing and what seem to be its most serious limitations?

In the text "Semantic Engines – An Introduction to Mind Design", which is the foreword to the book "Mind Design I", John Haugeland sketches the concept of Cognitive Science. He introduces the philosophical history, does a lot of description of terms that Cognitive Scientists use and introduces some key problems that Cognitive Scientists face.

<u>A short summary</u>
Haugeland begins with a little philosophical history that reaches back 300 years to thinkers like Hobbes and Descartes. Since then, he claims, three philosophical dilemmas have troubled a lot of heads. He mentions the debate of dualism versus materialism, the interaction between rules of thought and the rules of matter and the debate of mentalists versus behaviorists. He introduces the computational theory of the mind as the most promising theory to resolve these issues and as the inspirating thought of Cognitive Science.
He then goes on to introduce technical terms to define what a computer is. The first thing to say is that a computer is a Formal System. He describes Formal Systems as systems with a finite number of states. From this follows that they are self-contained, definite and digital.
The next step are Automatic Formal Systems. They are Formal Systems that perform steps on their own. Haugeland describes the famous Turing Machine and the thesis of Alonzo Church: any Automatic Formal System can be formally imitated by some Turing Machine. Special Turing Machines, called Universal Turing Machines, can even simulate every other Turing Machine (this was proven by Turing himself). From this follows that one is imposed to think about levels of abstractions that can make up a system: several layers of Turing Machines, imitating more and more abstract models on higher levels.
The discussion now leads to the Control Problem: Given that more than one next step is appropriate, how is decided which one is done? Haugeland describes how there are two "sub-machines": one to generate possibilies and another to chose from them. The bigger the complexity grows, the more difficult this chosing becomes. Relevant moves are therefore chosen in heuristic processes. What is often regarded as a "computer mistake" is often just due to very high complexity or bad implemented heuristics.
Haugeland then compares digital to analog systems. Analog systems sure have an infinite number of states and are useful in many circumstances, but digital systems are versatile and reduce the cost of errors. This makes a system affordable where an analog one just wouldn't be. Plus, digital systems can simulate analog systems.
When a system is to interact with the real world, one has to think about semantics. Tokens in the system get a meaning. This leads to an interpretation of the system. To chose among the possible interpretations is doing semantics. In a system with well-defined rules, the semantics take care of themselves. Those systems are called "Semantic Engines" and are a basic concept of Cognitive Science.
Haugeland then discusses how truth works as an interpretation concept (mostly in mathematics/logic). Such an interpretation should -in a desired way- "make sense", to be the chosen one. The story of chosing an interpretation in the real world is way more difficult. The concept of "making sense" is fuzzier than in mathematics or logic. One has to consider topics like rationality, the input/output transducers of the system and conversational cooperation. It is still under discussion how alternatives to truth-preservingness are to be defined.

Why the idea is appealing
As mentioned in the summary, the idea of minds being computers is a serious candidate for resolving big philosophical and scientific dilemmas. They bring up such widely discussed names as Descartes for philosophy or Skinner for psychology. Cognitive Science proposes a way out, using the most successful sciences of the last century: informatics, physics, mathematics, logic, (neuro)biology. The successes made in these fields over the last 150 years or so are overwhelming. It seems natural to give them a chance to try themselves on the difficult grounds of the human mind. Of course, the successes of some of those sciences raised the questions by themselves, for example chess computers. No philosopher was needed to raise the question if they might actually think (that sure does not hold for the answer...). The people who are successful in those areas are also mostly prone to materialism (which might not only be the reason but also a result of their successes).
In Addition, building digital models of the mind is a hands-on way of working. It works bottom-up, so you always see results, be they little. This learning by doing approach is new to sciences dealing with the mind (let aside brain surgery). An understanding of ourselves that works with digital models is also retractable and provable.
And as Haugeland also points out, digital means effective. On the long journey to find out if the mind is a computer, a lot of useful applications will come up on the way that might make our lifes easier.

Limitations to the idea
Haugeland adressed two philosophical critics: the "hollow shell" and the "poor substitute". The first claims that there is no real “x” inside, where “x” could be replaced with consciousness, original intentionality or caring. He claims that we need further criterions to answer this question. The second simply says that the computational approaches will never lead to useful results. This one will only be resolved if results come in one day.
Further problems are the algorithmical limitations of unsolvable problems and the complexity of pragmatics like the fuzzyness of “making sense” or the appropriateness in dealing with other machines or even humans. As we learned, faster or bigger computers will not help.

Another question that comes to mind is: If we use the state-of-the-art sciences of our time to resolve this holy grail of the sciences (the human brain), what would they use in another 200 years? Maybe we just don't know today what it takes to make it. A nice analogy would be the physician Galvani. When he made a dead frog's leg move, he thought to have discovered some electricity special to living things (Volta later discovered that there is no special electricity inside of living things). In those days (the 18^{th} and 19^{th} century) electricity was a widely disussed candidate for the secret ingredient of living things. It is therefore no wonder that science fiction about artificial life from those times like Mary Shelly's Frankenstein just puts electricity to dead bodys to make them live. Today, we know that it just takes basic mechanics and electronics to makes things move – but to think and feel is still a miracle to us. The science fiction of our days introduces artificial life in computers. What makes them “live” is their artificial personality, based on complex computation. I could imagine a future where sciences have further advanced such that people know for sure that being alive is not all about information processing and our approach has been too limited. Maybe they will know what it means to have consciousness (and to have none).
“Man's mind once stretched by a new idea, never regains its original dimension.” (Oliver Wendell Holmes)