Article

Intelligence without Representation: A historical perspective

Anna Jordanous

School of Computing, University of Kent, Chatham Maritime, Kent ME4 4AG, UK, a.k.jordanous@kent.ac.uk

Version September 1, 2020 submitted to Systems

Abstract: This paper reflects on a seminal work in the history of AI and representation: Rodney Brooks' 1991 paper *Intelligence without Representation*. Brooks advocated the removal of explicit representations and engineered environments from the domain of his robotic intelligence experimentation, in favour of an evolutionary-inspired approach using layers of reactive behaviour that operated independently of each other. Brooks criticised the current progress in AI research and believed that removing complex representation from AI would help address problematic areas in modelling the mind. His belief was that we should develop artificial intelligence by being guided by evolutionary development of our own intelligence, and that his approach mirrored how our own intelligence functions. Thus the field of behaviour-based robotics emerged. This paper offers a historical analysis of Brooks' behaviour-based robotics approach and its impact in artificial intelligence and cognitive theory at the time, as well as in modern-day approaches to AI.

Keywords: behaviour-based robotics; representation; adaptive behaviour; evolutionary robotics

1. Introduction

In 1991 Rodney Brooks published the paper *Intelligence without Representation*: a seminal work in behaviour-based robotics [1]. This paper influenced many aspects of research into artificial intelligence and cognitive theory.

In *Intelligence without Representation* Brooks described his behaviour-based robotics approach to artificial intelligence. He highlighted a number of points that he considers fundamental in modelling intelligence. ~~These ideas were controversial and triggered much debate.~~ Brooks was advocating the removal of explicit representations and engineered environments from the domain of his robotic intelligence experimentation. This was in stark contrast to traditional approaches to AI, which Brooks attacked in this paper, as well as contrasting with many modern-day approaches to AI.

Whilst Brooks' views have inspired similar research by a number of his contemporaries, ~~and brought about the research area of behaviour-based robotics as well as inspiring other research areas,~~ his ideas were controversial and ~~he also has~~ many critics ~~who~~ do not accept the validity of what he proposes. Nevertheless Brooks has impacted research in artificial intelligence and cognitive science since 1991. He presents a distinct view on how we can use artificial intelligence techniques to model and understand our own intelligence.

2. The state of Artificial Intelligence research prior to 1991, as the context for *Intelligence without Representation*

The two prominent schools of thought in Artificial intelligence (AI) prior to the publication of *Intelligence without Representation* were (i) traditional AI (also referred to as symbolic AI, logical AI, *Good Old Fashioned AI* (GOFAI) or classical AI) and (ii) connectionism (also referred to as neural networks). Traditional AI was the most prevalent approach at the time of writing *Intelligence without Representation* [1].

Brooks' critical comments on the failures of AI research to date were mainly addressed at traditional AI, although his criticisms also included connectionism, which was coming back into active research after developments in the 1970s. Although there is little in common between these two approaches, ~~they are widely regarded as both being derived from the same original source of McCullogh and Pitts' work~~ this was not always the case; for example the seminal work by McCullogh and Pitts on a neuronal approach to logic-based AI ~~[reference updated]~~ [2] influenced both traditional AI and connectionism [3].

To understand the AI environment which contextualises [1]: the aim of traditional AI at this time was to demonstrate how the processing of information can be done in an intelligent manner; generally efforts are focused on carrying out specific tasks or solving problems in specialised domains. The information being processed was represented in such a way that computation can be performed to create a solution. This would be done by performing manipulation of some symbolic format of the problem domain, for example Hayes suggested the use of first-order logic [4]. This manipulation is referred to as the physical symbol system hypothesis [5], with the symbolic manipulation centrally controlled by some supervisory process. The manipulation of a representation of a problem domain for intelligence can be seen today in knowledge representation research examples for contemporary robotics such as *KnowRob* and *RoboEarth* (see [6]), as well as in areas such as data mining and statistical machine learning algorithms.

Connectionism, on the other hand, was more biologically-inspired, taking as its model what we know about the architecture of the human brain. The concept behind connectionism was that if we construct an inter-connected network of units similar to the neural units that we believe the brain is constructed of, then we can construct an intelligent system that functions as the brain does. One can trace modern day research in deep learning, neural networks and cognitive neuroscience back to connectionist roots.

The overreaching contemporary goal for artificial intelligence as viewed by many commentators was to replicate human intelligence in a computational system [5,7,8].[1] Brooks strongly criticised the general state of AI research at that time [1,8], commenting that research was not advancing towards this goal as it should, and was too specialised into sub-disciplines such as planning or logical reasoning.[2] He argued that our lack of sufficient understanding of human intelligence prevents us from being able to partition AI effectively into specialisms that could be recombined in the future:

> "No one talks about replicating the full gamut of human intelligence any more. Instead we see a retreat into specialized subproblems ... Amongst the dreamers still in the field of AI ... there is a feeling that one day all these pieces will all fall into place and we will see "truly" intelligent systems emerge.
>
> However I, and others, believe that human level intelligence is too complex and little understood to be correctly decomposed into the right subpieces at the moment." [1, p. 140]

He believed that traditional AI research efforts addresses the wrong type of intelligent behaviour and quoted evolutionary timescales to support this view. For example, it took 5000 years to develop the art of writing: only a tiny fraction of the 3.5 billion years it took to develop the essence of "being and reacting" [1]. Given this evolutionary precedent, Brooks believed that the proportion of research in specialised areas such as computational linguistics and natural language processing is disproportionately large and that we should instead be concentrating our efforts on modelling more simple intelligent behaviour.

A powerful analogy used in [1] is that of artificial flight. He hypothesised about a group of artificial flight researchers of the 1890s (before aeroplanes were invented). These researchers are

[1] See concluding remarks to see how this definition has evolved since.

[2] Certainly AI was not progressing as hoped for: Turing expected his test of intelligence [9] to have been passed by the end of the twentieth century; however at the time of writing the Turing test still has not been passed.

transported in time to travel on an aeroplane in the 1980s. Given this glimpse of how artificial flight is real and possible, on their return they work on replicating the aircraft of the 1980s, exactly as they have seen, but with only their own limited knowledge to guide them.

His point is that modern day AI researchers are similar to these artificial flight researchers from the 1890s, in that they are trying to reconstruct the example of their work in practice that they have seen working successfully, with insufficient basic knowledge about the underlying foundations of their work. In other words, AI researchers in 1991 are working without sufficient knowledge of how our intelligence is developed from basic levels (like that demonstrated by insects, for example).

There is some precursor to Brooks' views in the discussion of the successes and failures of AI. Dreyfus and Dreyfus constructed a critical review of early work in Artificial Intelligence, up until the time of writing [10]. They judged that traditional AI has been found wanting and has not made the achievements in explaining and modelling the mind that it should. Hubert Dreyfus had previously argued [11] that intelligence could not be generated by symbol manipulation and that AI research was weakened by its reliance on Newell and Simon's physical symbol system hypothesis and felt that this argument had not been countered yet.

> "The physical symbol system approach seems to be failing because it is simply false to assume that there must be a theory of every domain." [10, p. 330]

Drew McDermott, previously a key exponent of logical AI, also described how he was now drawn more to an approach which discards deduction as the sole implementation of the thinking processes [12]. As McDermott still made use of knowledge representation, his approach in [12] was not as extreme a departure from traditional AI as Brooks's was. However this public shift in views from a prominent traditional AI researcher was indicative of problems in the traditional paradigm, particularly as McDermott expresses concern that a large proportion of AI research thus far may be leading to results that are of no use. McDermott admitted in [12] that his own research using this new approach in had not been overly fruitful and that he had been forced to retreat from this new viewpoint in his practical work, but he expressed confidence that once given proper investigation, this new "territory" could provide AI research with interesting new results. His later research, however, moved towards Semantic Web ontology research: an area emphasising significant use of knowledge modelling and representation.

Still, prominent voices were criticising traditional symbolic AI as a means of explaining human intelligence sufficiently. Therefore when Brooks' views were published, although controversial, they had been given some foundation in philosophical AI literature.

It might be surprising to learn that Brooks' academic background half a decade prior to publishing *Intelligence without Representation* (and similar papers) was rooted in symbolic AI. He wrote several papers in the early 1980s on work in planning, representation-based models and symbolic reasoning and analysis.[3] Brooks discusses in a 1997 interview with the Edge Foundation how the complexity of the mathematical models he was creating in his earlier career led him to the conclusion that the symbolic approach was not the right way to explain how intelligence worked.

> "It just couldn't be right. We had a very complex mathematical approach, but that couldn't be what was going on with animals moving their limbs about. Look at an insect, it can fly around and navigate with just a hundred thousand neurons. It can't be doing this very complex symbolic mathematical computations. There must be something different going on." [14]

[3] It is interesting to note that although Brooks includes these publications in his list of publications on his academic profile [13], none of the symbolic AI papers are linked to full texts of the paper (though the full texts are accessible from other sources), whereas the vast majority of his papers post 1985 do have links. ~~Brooks is acknowledging the existence of those papers but is disassociating himself from the contents of his earlier publications~~

It was this realisation that caused his change in perspective. From the mid 1980s onwards Brooks worked on developing this line of thought. This was to lead to the publication of *Intelligence without Representation* in 1991.

3. Fundamental aspects of *Intelligence without Representation*

3.1. Discarding Representation in favour of physical embodiment in the real-world environment

As suggested by the title of *Intelligence without Representation*, the hypothesis Brooks presented in 1991 is that:

> "Representation is the wrong unit of abstraction in building the bulkiest parts of intelligent systems" [1, p.140]

Brooks advocated that explicit central representations, "representation of goals that some central (or distributed) process selects from" [1][p. 144], are unnecessary to intelligence. Indeed they may even form a barrier to the demonstration of the mind's intelligence: as computations on representational symbols become more complex they take longer to process, hence Brooks' robots performed well compared to other contemporary robots in dynamic environments, because they could react in real-time without needing lengthy computational time. Brooks found during his own robotics experimentation that it is "better to use the world as its own model" [1, p. 140].

The behaviour-based robotics approach (also referred to by Brooks as *Nouvelle AI* in Brooks [8] requires that intelligence is demonstrated through our actions and interactions with the world. Critical to this is that the environment within which the robot operates must be independent of the robot design. It must not be simplified or targeted towards assisting the robot in any way. Instead the robot must be able to perform appropriate active and reactive behaviour to any environment it is put in.

As demonstration of this principle, one of Brooks' more famous intelligent robots (which he refers to as "Creatures") was *Herbert*, which operated in the MIT labs with the specific task of picking up drink cans. The lab environment is ever-changing, particularly when new people come to watch the robots operate. So Herbert would constantly need to react to new elements around it, working in real-time. Herbert was equipped with a small number of sensors and a mechanical arm that could grasp objects. Simple behaviours such as the ability to 'wander' around were enabled using infra-red and laser sensors to navigate corridors and doors, to follow walls and avoid obstacles, *as they were encountered*. Herbert's sensors could also detect objects available to be grasped by its arm. [15].

Herbert was illustrative of the initial success of Brooks' approach. In comparison, existing traditional AI robots at that time relied on being located in a static environment, a space which they would probably *hold an internal representation of*. For example, one could imagine that for a 'GOFAI Herbert' to solve the specific task of picking up drink cans in the MIT labs, it would require an internal representation of the MIT labs and objects in the labs. Movements would be calculated based on processing incoming input against that representation. Extra computation may be required on the occasions when input did not match what was expected given the robot's internal representations. Though this GOFAI Herbert is hypothetical rather than realised, generally in practice one could see how it might follow the pattern Brooks was observing: that compared to robots like Herbert, traditional AI robots ~~and~~ exhibited very little movement compared to processing time [14].

It is important to clarify what, in Brooks' eyes, constitutes a representation, so we understand precisely what Brooks is rejecting.[4] Brooks did not directly define the term 'representation' in *Intelligence without Representation*, but he did describe what he sees as a "good representation" [1, p. 140] for AI:

[4] As noted by a reviewer of this paper, "this failure to define the term "representation" fuels a lot of discussion in cognitive science. This implies that the term is not intuitively precise and that it means different things to different people. "

> "The idea was that by representing only the pertinent facts explicitly, the semantics of a world (which on the surface was quite complex) were reduced to a simple closed system once again. Abstraction to only the relevant details thus simplified the problems." [1, p. 140]

As an illustrative example, Brooks gave the following representation of a chair:

> "(CAN (SIT-ON PERSON CHAIR)), (CAN (STAND-ON PERSON CHAIR)) " [1, p. 140]

According to Brooks, a "standard representation" (i.e. a traditional AI representation) is composed of "variables ... that need instantiation in reasoning processes. ... rules which need to be selected through pattern matching. ... choices to be made" [1, p. 145]. Brooks rejected "traditional AI representations ... tokens which have any semantics that can be attached to them. " [1, p. 144].

Later, in [16], Brooks clarified his interpretation of representations as "an abstract manipulable entity" [p. 8], "[i]nternal world models which are complete representations of the external environment" [p. 3], that "rely on a semantic correspondence with symbols that the agent possesses" [p. 3]:

> "The common view in Artificial Intelligence, and particularly in the knowledge representation community, is that there is a central storage system which links together the information about concepts, individuals, categories, goals, intentions, desires, and whatever else might be needed by the system. In particular there is a tendency to believe that the knowledge is stored in a way that is independent from the way or circumstances in which it was acquired." [16, p. 14]
>
> "Over the years within traditional Artificial Intelligence, it has become accepted that they will need an objective model of the world with individuated entities, tracked and identified over time—the models of knowledge representation that have been developed expect and require such a one-to-one correspondence between the world and the agent's representation of it." [16, p. 16]

Brooks also used [16] to restate more clearly his position regarding representations, making it clear that he was not advocating for the entire removal of all representations of the environment that a robot operates in:

> "My earlier paper [*Intelligence without Representation*] is often criticized for advocating absolutely no representation of the world within a behavior-based robot. This criticism is invalid. I make it clear in the paper that I reject traditional Artificial Intelligence representation schemes (see section 5). I also made it clear that I reject explicit representations of goals within the machine. There can, however, be representations which are partial models of the world" [16, p. 21]

3.2. Subsumption architecture and emergent behaviour

The *subsumption architecture* reported by [1] for Brooks' robots is a modular structure composed of layers of simple behaviour such as *WANDER* or *AVOID OBSTACLES*. These layers co-exist independently and interact only as a side effect of their co-existence, rather than directly intending to communicate between layers. There is no central control of the layers or any symbol passing between layers; they operate independently of one another.

Emergentism, or emergent behaviour, is key to how the layers of Brooks' subsumption architecture operate in parallel to demonstrate intelligent behaviour. Brooks took Herb Simon's observation [17] of an ant walking across sand as an example of complex behaviour emerging from the combination of simple behaviours of an ant's movement reacting to a complex environment. Brooks suggested complex behaviour in intelligent creatures is more a result of complexity in the environment rather than in the intelligent creature. So complex behavioural patterns can emerge when simple behaviours combine together and are situated in an environment.

The incremental nature of Brooks' robots means that successful behaviours are retained in future. New robots are developments of older robots: they have the same behaviour of older robots except that they have additional layers in their subsumption architecture (this is instead of constructing a completely new architecture for each new robot).

Brooks attributed much of his inspiration to evolutionary precedents. Evolution involves a great deal of 'trial and error', where organisms with the more successful developments flourished whilst other organisms struggled to survive. In past reflections [18], Brooks has described how his research methods take a similar direction:

> "We don't have any plans! [for our robots] These are all research robots built experimentally ... patch upon patch upon kludge."

The successful 'patches' and 'kludges' survive and are retained, with new 'patches' added as the research develops.

4. Explaining cognitive processes through Brooks' approach: How does this differ to other AI approaches?

> "AI can help us understand the mind - what it is, as well as how it works." [7]
>
> "Computational studies, investigating similar problems being solved in a mind-body complex, help develop a more rigorous model and, more importantly, provide an understanding of the flow of information in and between these processes." [19]

The above two quotes are indicative of the commonly held view that AI research is a valuable tool in explaining how the mind works. Andy Clark describes how our mental processes are envisaged by proponents of traditional AI and of connectionism:

> "Classicists believe that thinking just is the manipulation of items having propositional or logical form; connectionists insist that this is just the icing on the cake and that *thinking* ('deep' thinking, rather than just sentence rehearsal) depends on the manipulation of quite different structure. As a result, the classicist attempts to give a level 2 processing model which is defined over the very same kinds of structure as figure in her level 1 theory.[5] Whereas the connectionist insists on dissolving that structure and replacing it with something quite different.
>
> A curious irony emerges. In the early days of Artificial Intelligence, the rallying cry was 'Computers do not crunch numbers, they manipulate symbols'. This was meant to inspire a doubting public by showing how much computation was like thinking. Now the wheel has come full circle. The virtue of connectionist systems, it seems, is that 'they do not manipulate symbols, they crunch numbers'. And nowadays we all know (don't we?) that thinking is not mere symbol manipulation! So the wheel turns." [21, p. 306]

Brooks, on the other hand, believes that thinking is an emergent phenomenon that arises as a result of different behavioural layers interacting with each other. His approach could be said to form a new Kuhnian paradigm in artificial intelligence and cognitive science. He treated traditional AI and connectionism as the *normal science* within which researchers are encountering difficulties such as the symbol-grounding problem or the frame problem. Behaviour-based robotics is the *revolutionary science*

[5] This is a reference to Marr's 3-level cognitive model of information processing [20]. Level 1, the *computational level*, looks at what a system does, what functions it performs, and why. Level 2, the *algorithmic level*, looks at how a system operates, and the representations and processes it employs. Level 3, not mentioned here, looks at the physical realisation of the system, and is perhaps the most relatable to Brooks' ideas (though by no means an accurate one-to-one mapping). Clark is making the point that, in contrast to a connectionist's perspective, someone from a classicist (traditional AI) perspective treats *how* we think (level 2) in terms of *what functions* may be occurring (level 1).

that Brooks presented as a new and more accurate paradigm that AI research needs to shift to in order to make any real further progress [22].

It is worth looking at Brooks' arguments on how behaviour-based robotics deals with the symbol-grounding problem and the frame problem. In ~~his~~ the Chinese Room analogy [23], Searle described how intelligence cannot be attained purely through manipulating symbols, (the prevalent attitude in symbolic AI). Instead these symbols should be "grounded" in some way so that they are given semantic validity. It is only when there is some meaning to the symbols that guides the manipulation of the symbols, that any such process could be considered as showing intelligent behaviour.[6]

This *symbol grounding problem* - defining the actual semantics or referential meaning of the computation performed by intelligent systems - was of major concern to symbolic AI practitioners. To Brooks and his supporters, however, this problem is essentially solved by behaviour-based robotics as a matter of course, by basing the robots in a real-world environment and using the surrounding environment to guide its actions rather than representations of the environment. Brooks believes that the symbol-grounding issue does in fact demonstrate how his behaviour-based approach addresses one of the fundamental weaknesses of traditional AI and connectionism [1,8].

Another classical problem in symbolic AI is known as the *frame problem*. If the world that the intelligent system operates in is given some form of representation, then the intelligent system has to deal with monitoring changes in the environment and incorporating these in the representation. Also any aspects of the environment that have not been encoded representationally may cause the intelligent system problems. As Brooks' earliest robots have shown, behaviour-based robotics encounter little difficulty when dealing with dynamic environments because they are embodied within the real world and interact with the world in real-time rather than abstracting a model representation of the world.

For any intelligent system to be deemed a proper model of the mind, it is necessary to define the criteria by which the system is being judged. A fundamental tenet of Brooks' theory is that the workings of the mind are demonstrated through intelligent behaviour. Hence the behaviour of his robots is indicative of the workings of an artificial mind.

Another key point for Brooks is that the human mind is not the only source of intelligence; other organisms also demonstrate intelligence which is worth modelling. He is very clear on this point. Human cognition can be extremely complex. Brooks described human-level cognition as "the holy grail of AI" and concedes that as of 1990:

> "neither classical nor nouvelle [behaviour-based] AI seem close to revealing the secrets of the holy grail of AI, namely general purpose human level intelligence equivalence" [8, p. 4]

It is plain that Brooks believes his Nouvelle AI to be the correct way in which to achieve this goal.

As an anti-representationalist, Brooks believes we do not need to rely on representation, instead we should gradually develop our intelligent systems, one step at a time. If the human mind developed through an incremental process such as this, then there is no wisdom in pursuing methods which require fresh representations to be constructed at each step [1,8,14].

Brooks also had issues with the level of complexity necessary to model some facets of the mind:

> "Symbol systems in their purest forms assume a knowable objective truth. It is only with much complexity that modal logics, or non-monotonic logics, can be built which better enable a system to have beliefs gleaned from partial views of a chaotic world.
>
> As these enhancements are made, the realization of computations based on these formal systems becomes more and more biologically implausible. But once the commitment to

[6] And even then, Searle says this is only Weak AI, or the simulation of intelligence, rather than the demonstration of true intelligence itself.

> symbol systems has been made it is imperative to push on through more and more complex and cumbersome systems in pursuit of objectivity." [8, p. 4]

Brooks' solution to this problem of complication and complexity was to eradicate the symbol systems and embrace a more biologically inspired, more simple methodology. His theories that there is no place for mental representations while performing intelligent behaviour have been preceded by similar observations. For example Dreyfus and Dreyfus described Wittgenstein's statements that our mental systems do not have entities within them that correspond to a singular thought or idea [10], and Gibson made the case that perception has no use for mental representations [24]. More recently, Harvey has warned of the need for 'linguistic hygiene' when using terms such as 'representation', arguing that while representations can be usefully studied in AI and cognitive science, there is much to be gained from a 'minimal cognition' approach which builds only the minimal model of cognition necessary to achieve the desired behaviour [25].

There is further criticism of the application of traditional AI as a model of our own intelligence. This criticism is based around the level of assistance traditional AI systems could be said to receive. Brooks believed that as of 1991 our efforts at building intelligent systems were misguided, in that we were providing the intelligent systems with too much of our own intelligence as assistance instead of getting the systems to demonstrate true intelligence.

> "Under the current scheme the abstraction is done by the researchers leaving little for the AI programs to do but search." [1, p. 143]

His point here was that the truly intelligent tasks are not undertaken and that researchers assist their systems too much; this is help which our own intelligence systems did not have when developing. Hence Brooks was very critical of the performance of existing AI methods thus far; artificial intelligence, he argued, was not true intelligence at all, but merely "simple numeric computations carried out in the sea of symbols". [8, p. 5]

A key point that Brooks emphasised[7] was the generality of intelligent behaviour. Our minds do not just focus on one task at a time. Even if our attention is drawn to one task that we are concentrating on, our minds are concurrently managing several other tasks at once. For example if a person is concentrating on reading a book, their mind will be simultaneously managing many other tasks such as holding the book, sitting in a chair, comprehending any conversation that is made to that person, and so on.

Brooks' strong criticism of traditional AI is that there is too much focus on building systems which are specialised for particular tasks or functions.

> "In classical AI, none of the modules themselves generate the behavior of the total system. Indeed it is necessary to combine together many of the modules to get any behavior at all from the system. " [8, p. 3]

He went on to ~~say~~ present his opinion that such a combination of classical AI modules is likely to present some difficulty because of the specialised way in which they have been designed [1].

Traditional AI and connectionism have been much maligned for having shortcomings, however they are not totally without benefit. Boden considered the key advantages of symbolic AI to be its representational ordering structure and clear definitions of problems to be solved and associated constraints [26]. In contrast, she saw the benefits of using a connectionist approach to be their biologically more plausible basis, trainability and 'gradual degradation' (should parts of the system stop working, the whole system does not suffer: the performance is reduced rather than completely halted).

But which approach is superior? It is time to take a closer look at Brooks' theory.

[7] And still emphasises - see the discussions later in this paper on Brooks' impact on and opinions of modern-day AI research.

5. Discussion of Brooks' approach

Hayes et al. raised some interesting counter-arguments to Brooks' views [27]. They made no secret of their belief that representations are very useful for modelling mental processes and should not be discarded. They used a 'nannies and babies' metaphor throughout their article to illustrate this. They did not deny that intelligent agents should be able to operate in and react to a social environment (situated AI), but they made a very strong claim about the nature of agents that do so without using some form of representation of that environment, concluding that such an approach may even cast doubts on the very validity of scientific method:

> "This perspective has its intellectual roots in parts of recent sociological thinking which reject the entire fabric of western science." [27]

The authors of [27] recognised that traditional AI does show flaws that need to be addressed, and agreed that there is a need for getting the basic foundations of cognitive theory correct. (One of the authors, Patrick Hayes, co-authored work with John McCarthy on philosophical problems within AI, including an early identification of the frame problem [28]. For people such as Brooks, though, who suggest modelling intelligence from a completely different theoretical perspective in order to solve these problems, Hayes et al. were critical, using the adage "Don't throw the baby out with the bathwater" [27]. They accused such people of ignoring crucial developments in modelling the mind (giving planning as an example), and of condemning these developments as unworthy of retention, in their haste to follow new cognitive theories. This was an emphatic attack on Brooks and his contemporaries, but perhaps it makes more extreme conclusions than is deserved, in its attempt to provoke reaction.

Some have questioned whether behaviour-based robots do actually rely on representations in an inferential form without being explicitly coded (for example [29]). Brooks anticipated such suggestions and dismissed them with the comments that:

> "There are no variables ... that need instantiation in reasoning processes. There are no rules which need to be selected through pattern matching. There are no choices to be made. To a large extent the state of the world determines the action of the Creature" [1, p. 149]

So Brooks firmly advocated that he has not used any form of representation-based methods. This is however still open to much debate, as has been seen in the literature (for example [29,30]).

It must be questioned whether symbolic representation is as limited as Brooks describes. Brooks does have experience in working in symbolic AI prior to his shift in thinking, and described many problems that he ran into that he described as a direct consequence of using representation to explain how the mind works. However it is generally conceded, as a result of [23] and similar writing (for example [27, pp. 17-20], that intelligence is not purely just about the rigid manipulation of symbols, but requires some guidance from knowledge of the semantics behind the symbols. Some forms of intelligence modelling do have a more simplified use of knowledge representation (in fact Hayes et al. imply that these include the research that Brooks was part of in the early 1980s [27, p. 17]). Proponents of symbolic AI argued that knowledge representation is far more flexible and less restricted than such simplified systems would suggest. For example, advances have been made employing representation learning, such as in learning object affordances [31,32] or state information [33]. This debate on the validity of symbolic representation of knowledge for cognitive purposes has been in progress since the publication of Brooks' ideas and, with current interests in representations for robotics (e.g. [6]) as well as more generally in areas such as data mining, does not look close to resolution.

Whilst discussing cognitive science models that are embodied, having physical presence, [34], Andy Clark made a valid observation concerning Brooks' assertion that a behaviour-based robot shows intelligent behaviour if it carries out some reactive behaviour that is helpful to the survival of the robot. Clark described how a sunflower will react to the changes in positioning of the sun by changing the direction in which it faces, in order to maximise the exposure to the sun it receives [34, p. 347]. The question is whether this is a demonstration of intelligent cognition and behaviour by the sunflower. The natural reaction to this is that a sunflower does not carry out any rational thought process, so it cannot be demonstrating intelligence. However by strict interpretation of Brooks' assertions, this sunflower would be considered intelligent. Clearly this aspect of Brooks' theory could benefit from clarification, particularly on what constitutes intelligent behaviour.[8]

In his critical response to [1], Etzioni made some interesting points in direct reference to the evolutionary influence that Brooks gives as justification for his methods [35]. Etzioni quoted the theory of punctuated equilibria which states that evolutionary development is not necessarily linearly proportional to the time taken but may include a high amount of variance. There may be little or no progress for a large amount of time, followed by a rush of new breakthroughs. Etzioni used this to questions Brooks' basis in evolutionary development of intelligence. Essentially he was asking if Brooks is selecting only the aspects of evolution that are useful for his argument, as for the example above with the comparison of timescales of different evolutionary developments. Etzioni asked exactly how Brooks has derived the conclusion that higher level cognition will follow on naturally from simple reactive robotics. He suggested that an equally valid conclusion could be drawn in the same way as Brooks' conclusion, that if AI researchers work towards developing the hardware of actual organisms, that the intelligent aspects of the mind will follow naturally [35, p. 9]. It is an interesting observation and not one that Brooks has chosen to answer, as far as I have been able to find.

The evolutionary justification used by Brooks also does not fully account for other decisions Brooks made within his approach. Behaviour-based robotics emphasises the building of simple robots at first, only adding complexity when the simpler robots have been successfully achieved. Brooks advocated this in [1], but argued that it is not right to have the same approach to the robots' environments (the use of simple environments at first, with complexity added incrementally). He justified this with the argument that evolution has proven the first approach to be successful, but that the second approach could mean that errors were inadvertently introduced [1] (in particular p. 150). I see, however, two problems with this argument. Firstly, as Brooks himself acknowledges:

> "As a very approximate hand waving model of evolution, things get built up and accreted over time, and maybe new accretions interfere with the lower levels." [14]

So Brooks concedes that his methods introduce unexpected behaviour which may not be a desired result (as is the nature of systems that make use of emergentism). Even if not regarded as errors, these unexpected behaviours are still a source of uncertainty.[9]

Also, Brooks did not consider that our environments have developed in complexity as we have evolved. As the human race[10] has evolved, our environments have received increasingly more complex adaptations as we become more technologically advanced: from the invention of the wheel to modern day transport systems and beyond. Also, as infants our environments are necessarily restricted by our parents or carers. Complexity is added as we are gradually exposed to more of the world. It is not considered plausible that a baby would be able to cope with a real-world environment in the way that a grown adult has learnt to; their world is controlled and simplified for them. It is to be supposed that Neolithic man in the modern-day world would have similar difficulties to the baby. However Brooks disregarded these aspects of human development and evolution in [1].

The emergent behaviour demonstrated by Brooks' robots was shown to be successful for relatively small tasks such as collection of drinks cans or map navigation - lower-level cognitive processes [36]. A common comment, however, is that designing a task-driven system which produces emergent

[8] But Brooks' emphasis is on research of a implementional nature rather than theorising so it is unlikely he would enter into much philosophical debate about his approach. Instead Brooks offers his robots as direct evidence of his theory.

[9] Uncertainty may, of course, be no bad thing.

[10] As a reviewer of this article notes, arguably this argument could be extended to animals as well.

behaviour is a hard task due to the element of uncertainty in prediction of emergent behaviour (for example [30], or even Brooks himself [37])~~.~~ ; rather than designing *for* a particular desired behaviour, an emergent system's behaviour develops over time. Following on from Brooks' own criticisms that we do not understand intelligence well enough to subdivide it into specialised sub-problems, a similar criticism could be levelled at Brooks' subsumption architecture design: do we know enough about emergentism to develop the correct layers of behaviour such that some required behaviour will emerge? And how do we know we have included all necessary layers in our design, if we cannot identify what these layers should be?

It is doubtful that Brooks would attach too much significance to these criticisms, from his previous comments:

> "Nouvelle AI relies on the emergence of more global behavior from the interaction of smaller behavioral units. As with heuristics there is no a priori guarantee that this will always work. However, careful design of the simple behaviors and their interactions can often produce systems with useful and interesting emergent properties." [1]

Indeed, to respond to these criticisms, one might point out that behaviour-based robotics is an evolutionary and incremental process of development where each new layer of behaviour gives the robot more complex intelligent behaviour. Rather than aiming for a target level of intelligent behaviour for a given task, we should be developing the robots' intelligence in a step-by-step manner, gradually increasing in complexity and generality. This is as inspired by evolution: the development of intelligence should be guiding our investigations into the modelling of the mind, not vice. versa. Questions still remain, though, as to whether AI research should take the same path as evolution (Brooks believes so; many critics disagree) and whether we have the epistemological knowledge necessary to learn from evolution in developing artificial intelligence.

A major test for behaviour-based robotics could be how it can be used to demonstrate wider or more general intelligence, particularly in comparison to more traditional methods. Brooks pointed out in 1991 that, as for traditional methods, behaviour-based AI should be allowed time to develop. On the criticism of behaviour-based robotics on the grounds that it cannot solve all tasks considered, he likens this to saying that an elephant has no cognitive processes worthy of study because it is unable to play chess[11] [8]. After nearly thirty years of research, however, the behaviour-based approach shows no signs of becoming the dominant paradigm in AI research. This is not to say that it is impossible for behaviour-based robotics to reach their intended goal of developing highly intelligent robots; but to date we have not yet seen the full scope of Brooks' visions for AI robotics being realised.

The issue of scaling up to larger domains and more complex behaviour did in fact lead Brooks to revise his ideas in some ways [26]. For example he has had to relax his emphasis on pure reactive behaviour and allow that keeping some memory of previous experiences is necessary for higher level cognition. For example, his robot *Toto* learns about its environment by building maps of the parts of the environment already encountered in exploration. Brooks maintained that this is not a representation of the environment as it does not try to replicate the environment internally but instead recalls the sonar readings and actions made by the robot to manoeuvre around obstacles and walls. As said above, though, this argument is open to some debate.

[11] Brooks refers to the inability of elephants to understand the game of chess rather than the rather obvious physical difficulties they would have in picking up the pieces.

6. Brooks' impact and views on subsequent artificial intelligence research

6.1. Behaviour-based robotics after Intelligence Without Representation

The key area influenced by Brooks' ideas has been in robotics, where behaviour-based robotics is now an accepted component of robotic architecture [14,29,30,38,39]. One key example of the application of behaviour-based robotics is in Nasa's Mars Exploration Rovers [40], robots which operate autonomously on the surface of Mars, using behaviour-based principles. Behaviour-based robotics research such as [41–43] forms a small but recognised part of the broader field of evolutionary robotics [44]. It also sits within Artificial Life and adaptive behaviour research [45–47, e.g.]. Brooks continues to contribute to these areas [48, e.g.], though he has recently remarked that their progress has 'stalled' [49].

Brooks-style ideas have also had wider impact. Alan Bundy is a strong logical AI exponent, however he and Fiona McNeil have written that:

> "automatic representation development, evolution, and repair must be a major goal of AI research over the next 50 years." [50, p. 85]
>
> [and]
>
> "Reasoning systems must be able to develop, evolve, and repair their underlying representations as well as reason with them. The world changes too fast and too radically to rely on humans to patch the representations." [50, p. 86]

In other words, there is some acknowledgement that syntax, semantics and pragmatics of representations should be ~~evolved~~ discoverable rather than explicitly coded. This is happening to some extent in current robotics research such as [33], as well as more broadly in work applying evolutionary computation approaches, such as in computational creativity applications [51]. Such advances are exemplar of the far-reaching impact of evolutionary or biologically-inspired modelling of the mind, and the acknowledgement of the need to consider problems with traditional AI's use of representations.

Etzioni suggested in his direct response to *Intelligence without Representation* that Brooks's proposals would also work equally well for "soft-bots" in a real-time software environment (designed externally and outside of the control of the soft-bots' designers: for example the World Wide Web) [35]. In other words, physical embodiment is not necessary for demonstrating intelligent cognitive processes. Although Brooks has been dismissive of soft-bots in favour of "physical robots made of metal" [14], it is uncontroversial to assume that virtual agents can demonstrate intelligence in a virtual environment just as physical robots do in a real-world environment [52,53, e.g.].

Looking back at further direct responses to [1], Nilsson included Brooks' work in his review of AI research [36]. He presented a similar thesis to Brooks, that AI research has become over-specialised, and praised Brooks' work as a step in the right direction but criticised Brooks' approach as being overly focused on low-level processes, acknowledging the existence of doubts about "the long-range potential for this work"[36, p. 14].

Nilsson's conclusion was that research activity in AI to date (including Brooks' work) should be considered merely as development of tools to be used in the future by systems that display more general intelligence. Such hybrid approaches to AI have produced strong results in robotics research. For example Brooks describes how the Mars exploration unit Pathfinder takes Brooks' architecture as low level processing of information, with a classical AI-based representation of cognition at a higher level of processing. This layered model bears strong resemblance to the general perception of our minds as having higher level cognitive processes such as for playing chess, and lower level processes such as reacting to unstable ground when walking. The successes demonstrated with this approach are at least in part due to how different approaches to cognition apply on wider scales:

> "[The hope of Traditional AI] is that the ideas used will generalize to robust behavior in more complex domains. ... [The hope of Nouvelle AI] is that the ideas used will generalize to more sophisticated tasks.
>
> Thus the two approaches appear somewhat complementary. It is worth addressing the question of whether more power may be gotten by combining the two approaches."[12] [8]

Clark made a similar point:

> "As tasks become more representation-hungry - more concerned with the distal, abstract and non-existent - we will see more and more evidence of *some kinds* of internal representation and inner models." [34, p. 349]

Brooks has inspired research both within the field of robotics [30, e.g.] and more broadly, within cognitive theory [54]. Brooks' theories have contributed to research areas such as [26] Artificial Life (where emergent behaviour is a fundamental demonstration of intelligence) [25, e.g.] and adaptive behaviour approaches [55, e.g.]. His ideas sat alongside related contemporary research movements such as the *animat* approach (where the initial intelligent system start with simple behaviour and gradually the complexity of intelligence is built up, mimicking how life on Earth has evolved) [56], and Braitenburg's *Vehicles* [57], which similarly evolved complex behaviour from simple principles and has inspired fields such as swarm intelligence [58, e.g.]. Overall, the growth of evolutionary and adaptive approaches to AI e.g. [44,45,59, e.g.] has contributed to a broader perspective on AI beyond traditional AI and connectionism.

Brooks' advancement in academic circles to the post of Director of the Computer Science and Artificial Intelligence Laboratory at MIT shows how his work was taken seriously in academia. He has received various accolades, including the prestigious *IJCAI Computers and Thought Award* in 1991, presented to "outstanding young scientists in artificial intelligence".[13] The title of this award is interesting given Brooks' critique of the metaphor of computation in cognition. In response, Brooks reflected [16] that "Computers and Thought are the two categories that together define Artificial Intelligence as a discipline. It is generally accepted that work in Artificial Intelligence over the last thirty years has had a strong influence on aspects of computer architectures. In this paper we also make the converse claim; that the state of computer architecture has been a strong influence on our models of thought."

His 2002 book *Flash and Machines* [60], aimed at non-academics as well as academics, shows the extent to which Brooks' behaviour-based research has progressed his views. In this book, Brooks stated that humans are intelligent machines (biological machines, but machines nonetheless):

> "I believe myself and my children all to be mere machines. But this is not how I treat them. I treat them in a very special way, and I interact with them on an entirely different level ... I maintain two sets of inconsistent beliefs and act on each of them in different circumstances. It is this transcendence between belief systems that I think will be what enables mankind to ultimately accept robots as emotional machines."[60]

To date, Brooks continues to advocate behaviour-based robotics and the principles behind *Intelligence Without Representation*. In 2008 he moved to industry, starting up Rethink Robotics [14] as CTO, while still retaining some presence in academia as emeritus professor at MIT. He was part of the panel behind the 2016 'One Hundred Year Study on Artificial Intelligence (AI100)' Stanford report

---

12 Although Brooks later goes on to say that it is best to use his approach on its own rather than in combination with other AI approaches [14]

13 https://www.ijcai.org/awards, last accessed June 2020.

14 http://rethink.ai

[61]. At Rethink Robotics, Brooks has recorded several patents since 2012 that employ behaviour-based robotics commercially.[15]

6.2. Brooks' views on other areas in current AI research

What of Brooks' views comments on other areas of AI research today?

Brooks has expressed vocal opinions on various other topics in modern day AI research. For example, one way in which autonomous robotics research has entered modern-day consciousness with the move towards self-driving cars. While he has confidence in the technicalities of self-driving cars being realised, Brooks has voiced concerns about how people's behaviour will have negative effects on the pace of technical development. In an ironic twist, he sees that human behaviour will evolve to hinder robotic developments [37].

Machine learning is a currently popular area of Artificial Intelligence that heavily relies upon statistical representations. Perhaps unsurprisingly, Brooks has shown negativity towards this style of approach, issuing a scathing attack in his blog:

> "In 1991 I wrote a long ... paper[16] on the history of Artificial Intelligence and how it had been shaped by certain key ideas. In the final paragraphs of that paper I lamented that there was a bandwagon effect in Artificial Intelligence Research, and said that "[m]any lines of research have become goals of pursuit in their own right, with little recall of the reasons for pursuing those lines".
>
> "I think we are in that same position today in regard to Machine Learning. The papers in conferences fall into two categories. One is mathematical results showing that yet another slight variation of a technique is optimal under some carefully constrained definition of optimality. A second type of paper takes a well know learning algorithm, and some new problem area, designs the mapping from the problem to a data representation ... and show the results of how well that problem area can be learned.
>
> "This would all be admirable if our Machine Learning ecosystem covered even a tiny portion of the capabilities of human learning. It does not. And, I see no alternate evidence of admirability.
>
> "Instead I see a bandwagon today, where vast numbers of new recruits to AI/ML have jumped aboard after recent successes of Machine Learning, and are running with particular versions of it as fast as they can. They have neither any understanding of how their tiny little narrow technical field fits into a bigger picture of intelligent systems, nor do they care. They think that the current little hype niche is all that matters, are blind to its limitations, and are uninterested in deeper questions." [Postscript to [65]]

Brooks has also expressed scepticism over deep learning, an area of Artificial Intelligence currently strongly in favour and gaining large traction. In a 2012 Nature comment, Brooks warned that:

> "we are in an intellectual cul-de-sac, in which we model brains and computers on each other, and so prevent ourselves from having deep insights that would come with new models." [66]

This reads partly as an unspoken attack on the then-emerging area of deep learning. Deep learning relies heavily on multi-level ('deep') representations, often based around neural networks.

Brooks is not alone in his concern over the specificity of current machine-learning-focussed approaches to AI, and a lack of adaptability in AI applications to date. For example, as expressed in 2019 by Judea Pearl:

---

[15] e.g. patents [62] granted 2019, [63] granted 2016, [64] granted 2015.

[16] The paper that Brooks refers here to is [16]

> "The dramatic success in machine learning has led to an explosion of artificial intelligence (AI) applications and increasing expectations for autonomous systems that exhibit human-level intelligence. These expectations have, however, met with fundamental obstacles that cut across many application areas. One such obstacle is adaptability, or robustness. Machine learning researchers have noted current systems lack the ability to recognize or react to new circumstances they have not been specifically programmed or trained for." [67]

Brooks invests more faith in artificial general intelligence (AGI), though he warns that in his opinion, progress in the area of AGI is lacking at present [68]. AGI research investigates how AI can operate with general intelligence that can apply across multiple domains or tasks, rather than intelligence focused on specific tasks or domains. This is certainly an area where one can see the principles behind *Intelligence without Representation* contributing [69, e.g.], though it is not the case that AGI thus far has required a behaviour-based approach [70,71, e.g.].

In 2018, Brooks made several predictions for various aspects of artificial intelligence [72]. These include several attacks on deep learning, concluding that by 2027 we will reach the end of "the era of Deep Learning" and the "[e]mergence of the generally agreed upon "next big thing" in AI beyond deep learning." In contrast, he sees developments in self-driving cars and other robotics continuing over future decades. He is appraising these predictions annually; only time will tell as to how correct his predictions are.

7. Concluding remarks

As noted above, the goal of artificial intelligence prior to *Intelligence without Representation* [1] was to replicate human intelligence in a computational system [5,7,8]. Taking a modern and widely cited definition of artificial intelligence, Russell and Norvig define AI as "the designing and building of intelligent agents that receive percepts from the environment and take actions that affect the environment" [73]. This definitional shift is a sign that, even if Brooks' ideas are not directly responsible for the change, they are part of a broadening in how artificial intelligence is conceptualised.

Brooks is an elegant writer with a clear and well-presented style, put forward in a persuasive manner. His views have gained significant traction. The evolution of the human mind has been taking place over an extremely large time-scale. Given our current lack of a complete and agreed understanding of what intelligence is and how our mind works, should we be aiming so high in building computerised intelligent systems to display human-level intelligence? ~~The principles of Occam's razor (where a parsimonious and simple approach is favoured where possible) is often of use in furthering academic progress. This has been demonstrated in countless examples, from the preference of Copernicus' simpler planetary model over the convoluted Ptolemaic system, to heuristics for good practice in medical diagnosis. So Brooks' simple, 'bottom-up' approach could be considered as the practice of good scientific method.~~ Perhaps a simpler approach is needed.

But has Brooks discovered a definitive method of explaining and modelling the mind? Brooks's work showed impressive results at first. But, as the sophistication of the robots has increased and their behaviour has become more complex, development has slowed in pace somewhat. But this is reminiscent of the pattern of development of traditional or 'good old fashioned' AI. As noted above, Brooks now predicts that the currently popular AI approach of deep learning will follow a similar pattern [72].

Perhaps Brooks himself best summarises the situation between the competing paradigms in AI. Although written in 1990, this observation still holds:

> "Can there be a theoretical analysis to decide whether one organization for intelligence is better than another? Perhaps, but I think we are so far away in understanding the correct way of formalizing the dynamics of interaction with the environment that no such theoretical results will be forthcoming in the near term." [8, p. 13]

All is not doom and gloom, though, for AI research:

> "We have only just begun to explore the space of computational possibilities [for modelling our minds]. Changes of direction, and even the occasional dead end, should not be scorned as folly. Science grows not only by conjectures, but also by refutations." [26, p. 9]

Funding: This research received no external funding.

Acknowledgments: I acknowledge gratefully the careful input and encouragement from Paola di Maio, as well as the useful comments offered by the anonymous reviewers during peer review.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Brooks, R.A. Intelligence without Representation. *Artificial Intelligence* **1991**, *47*, 139–159.
2. McCulloch, W.S.; Pitts, W.H. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics* **1943**, *7*, 115–133. Reprinted in [74], pp. 16-39.
3. Piccinini, G. The First Computational Theory of Mind and Brain: A Close Look at McCulloch and Pitts's "Logical Calculus of Ideas Immanent in Nervous Activity". *Synthese* **2004**, *141*, 175–215. doi:10.1023/B:SYNT.0000043018.52445.3e.
4. Hayes, P.J. The Naive Physics Manifesto. In *The Philosophy of Artificial Intelligence (1990)*; Edinburgh University Press: Edinburgh, 1979; pp. 171–205. First published in Expert Systems in the Micro- Electronic Age (1979).
5. Newell, A.; Simon, H.A. Computer Science as Empirical Inquiry: Symbols and Search. *Communications of the ACM* **1976**, *19*, 113–126. The Tenth ACM Turing Lecture.
6. Paulius, D.; Sun, Y. A Survey of Knowledge Representation in Service Robotics. *Robotics and Autonomous Systems* **2019**, *118*, 13 – 30. doi:https://doi.org/10.1016/j.robot.2019.03.005.
7. Boden, M.A., Ed. *The Philosophy of Artificial Intelligence*; Oxford University Press, 1990.
8. Brooks, R.A. Elephants Don't Play Chess. *Robotics and Autonomous Systems* **1990**, *6*, 3–14.
9. Turing, A.M. Computing Machinery and Intelligence. *Mind* **1950**, *LIX*, 433–460.
10. Dreyfus, H.L.; Dreyfus, S.E. Making a Mind versus Modelling the Brain: Artificial Intelligence Back at a Branch-Point. In *The Philosophy of Artificial Intelligence (1990)*; First published in Artificial Intelligence 117 No. 1, 1988; pp. 309–333.
11. Dreyfus, H.L. *What Computers Can't Do: The Limits of Artificial Intelligence*; Harper and Row, 1979.
12. McDermott, D. A Critique of Pure Reason. *Computational Intelligence* **1987**, *3*, 151–160.
13. Brooks, R.A. Rodney Brooks - Roboticist. https://people.csail.mit.edu/brooks/publications.html, last accessed June 2020, 2008.
14. Brooks, R.A.; Brockman, J. The Deep Question: A Talk with Rodney Brooks. https://www.edge.org/conversation/rodney_a_brooks-the-deep-question , last accessed June 2020, 1997.
15. Brooks, R.A.; Connell, J.; Ning, P. Herbert: A Second Generation Mobile Robot. AI Memos (1959-2004) AIM-1016, MIT, 1988.
16. Brooks, R.A. Intelligence without Reason. Technical Report [Computers and Thought, IJCAI-91], Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1991.
17. Simon, H.A. *The Sciences of the Artificial*; MIT Press, 1969.
18. Brooks, R.A. FAQ. http://people.csail.mit.edu/brooks/faq.shtml, now inaccessible but retrievable via https://web.archive.org/web/20070225132932/http://people.csail.mit.edu/brooks/faq.shtml , last accessed June 2020, 2007.
19. Yeap, W.K. Emperor AI, where is your new mind? *AI Magazine* **1997**, *Winter*.
20. Marr, D. Artificial intelligence: a personal view. In *Mind Design*; Haugeland, J., Ed.; MIT/Bradford Books (1981): Cambridge, Massachusetts, 1977.
21. Clark, A. Connectionism, Competence and Explanation. In *The Philosophy of Artificial Intelligence(1990)*; Also published in The British Journal for the Philosophy of Science, 1990; pp. 281–308.
22. Kuhn, T.S. *The Structure of Scientific Revolutions*; The University of Press: Chicago, 1962.

23. Searle, J.R. Minds, Brains and Programs. *Behavioral and Brain Sciences* **1980**, *3*, 417–424.
24. Gibson, J.J. *The Ecological Approach to Visual Perception*; Houghton-Mifflin, 1979.
25. Harvey, I. Misrepresentations. Proceedings of the Eleventh International Conference on Artificial Life; Bullock, S.; Noble, J.; Watson, R.A.; Bedau, M.A., Eds.; MIT Press, , 2008; pp. 227–233.
26. Boden, M.A. New breakthroughs or dead-ends? *Philosophical Transactions of Physical Sciences and Engineering* **1994**, *349*, 1–13.
27. Hayes, P.J.; Ford, K.M.; Agnew, N. On Babies and Bathwater: A Cautionary Tale. *AI Magazine* **1994**, *Fall*, 15–26.
28. McCarthy, J.; Hayes, P.J. Some Philosophical Problems from the Standpoint of Artificial Intelligence. In *Machine Intelligence 4*; Meltzer, B.; Michie, D., Eds.; Edinburgh University Press, 1969; pp. 463–502.
29. Avraham, H.; Chechik, G.; Ruppin, E. Are There Representations in Embodied Evolved Agents? Taking Measures. *Lecture notes in Artificial Intelligence* **2003**, *2801*, 743–752.
30. Steels, L. Intelligence with Representation. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* **2003**, *361*, 2381–2395.
31. Min, H.; Yi, C.; Luo, R.; Zhu, J.; Bi, S. Affordance research in developmental robotics: a survey. *IEEE Transactions on Cognitive and Developmental Systems* **2016**, *8*, 237–255. doi:10.1109/TCDS.2016.2614992.
32. Zech, P.; Haller, S.; Lakani, S.R.; Ridge, B.; Ugur, E.; Piater, J. Computational models of affordance in robotics: a taxonomy and systematic classification. *Adaptive Behavior* **2017**, *25*, 235–271. doi:10.1177/1059712317726357.
33. Jonschkowski, R.; Brock, O. Learning State Representations with Robotic Priors. *Autonomous Robots* **2015**, *39*, 407–428.
34. Clark, A. An embodied cognitive science? *Trends in Cognitive Sciences* **1999**, *3*, 345–351.
35. Etzioni, O. Intelligence without Robots: A Reply to Brooks. *AI Magazine* **1993**, *Winter*, 7–13.
36. Nilsson, N.J. Eye on the Prize. *AI Magazine* **1995**, *Summer*, 9–17.
37. Brooks, R. The big problem with self-driving cars is people. *IEEE Spectrum: Technology, Engineering, and Science News* **2017**.
38. Arkin, R.C. *Behavior-based robotics*; MIT press, 1998.
39. Michaud, F.; Nicolescu, M. Behavior-based systems. In *Springer Handbook of Robotics*; Siciliano, B.; Khatib, O., Eds.; Springer: Cham, 2016.
40. Watanabe, S.; Dunbar, B. People Are Robots, Too. Almost. https://www.nasa.gov/vision/universe/roboticexplorers/robots_like_people.html , last accessed June 2020, 2007.
41. Lyons, D.M.; Arkin, R.C.; Jiang, S.; Liu, T.M.; Nirmal, P. Performance verification for behavior-based robot missions. *IEEE Transactions on Robotics* **2015**, *31*, 619–636.
42. Martín, F.; Aguero, C.E.; Canas, J.M. A Simple, Efficient, and Scalable Behavior-Based Architecture for Robotic Applications. Robot 2015: Second Iberian Robotics Conference. Springer, 2016, pp. 611–622.
43. Lee, G.; Chwa, D. Decentralized behavior-based formation control of multiple robots considering obstacle avoidance. *Intelligent Service Robotics* **2018**, *11*, 127–138.
44. Nolfi, S.; Bongard, J.; Husbands, P.; Floreano, D. Evolutionary robotics. In *Springer Handbook of Robotics*; Springer, 2016; pp. 2035–2068.
45. Rajagopalan, P.; Holekamp, K.E.; Miikkulainen, R. Factors that Affect the Evolution of Complex Cooperative Behavior. ALIFE 2019: The 2019 Conference on Artificial Life; , 2019; pp. 333–340.
46. Urashima, H.; Wilson, S.P. A Self-organising Animat Body Map. Living Machines: Conference on Biomimetic and Biohybrid Systems; Springer International Publishing: Milan, Italy, 2014; pp. 439–441.
47. Williams, P.; Beer, R. Environmental Feedback Drives Multiple Behaviors from the Same Neural Circuit. ECAL 2013: The Twelfth European Conference on Artificial Life; , 2013; pp. 268–275.
48. Brooks, R. The Philosophical Underpinnings of Work in Artificial Life. ALIFE 2018: The 2018 Conference on Artificial Life; , 2018.
49. Brooks, R.A. The Seven Deadly Sins of Predicting the Future of AI. https://rodneybrooks.com/the-seven-deadly-sins-of-predicting-the-future-of-ai/ , last accessed June 2020, 2017.
50. Bundy, A.; McNeill, F. Representation as a Fluent: An AI Challenge for the Next Half Century. *IEEE Intelligent Systems* **2006**, *May/June*, 85–87.

51. Cunha, J.M.; Martins, P.; Lourenço, N.; Machado, P. Emojinating Co-Creativity: Integrating Self-Evaluation and Context-Adaptation. 11th International Conference on Computational Creativity; , 2020.
52. Dahl, S.; Cenek, M. Towards Emergent Design: Analysis, Fitness and Heterogeneity of Agent Based Models Using Geometry of Behavioral Spaces Framework. ALIFE 2016: The Fifteenth International Conference on the Synthesis and Simulation of Living Systems; , 2016; pp. 46–53.
53. Berners-Lee, T.; Hendler, J.; Lassila, O. The Semantic Web. *Scientific American* **2001**, *284*, 34–43.
54. Wallis, P. Intention without Representation. *Philosophical Psychology* **2004**, *4*, 209–223.
55. Konidaris, G.D.; Hayes, G.M. An architecture for Behavior-Based Reinforcement Learning. *Adaptive Behavior* **2005**, *13*, 5–32.
56. Wilson, S.W. The animat path to AI. From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior; Meyer, J.A.; Wilson, S.W., Eds. The MIT Press, 1991, pp. 15–21.
57. Braitenberg, V. *Vehicles: Experiments in synthetic psychology*; MIT press, 1986.
58. al Rifaie, M.; Bishop, J.; Caines, S. Creativity and Autonomy in Swarm Intelligence Systems. *Cognitive Computation* **2012**, *4*, 320–331.
59. Jordanous, A. A Fitness Function for Creativity in Jazz Improvisation and Beyond. International Conference on Computational Creativity; , 2010; pp. 223–227.
60. Brooks, R.A. *Flesh and Machines: How Robots Will Change Us*; Pantheon Books, 2002.
61. Stone, P.; Brooks, R.; Brynjolfsson, E.; Calo, R.; Etzioni, O.; Hager, G.; Hirschberg, J.; Kalyanakrishnan, S.; Kamar, E.; Kraus, S.; Leyton-Brown, K.; Parkes, D.; Press, W.; Saxenian, A.; Shah, J.; Tambe, M.; .; Teller, A. "Artificial Intelligence and Life in 2030." One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel. Available at http://ai100.stanford.edu/2016-report. Technical report, Stanford University, Stanford, CA, 2016.
62. Brooks, R.; Caine, M. Patent: Hybrid training with collaborative and conventional robots. https://patents.google.com/patent/US10514687B2/, last accessed June 2020, 2019.
63. Ivanov, Y.A.; Brooks, R. Patent: Robotic placement and manipulation with enhanced accuracy. https://patents.google.com/patent/US9457475B2/ , last accessed June 2020, 2016.
64. Brooks, R.; Buehler, C.J.; Cicco, M.D.; Ens, G.; Huang, A.; Siracusa, M.; Williamson, M.M. Patent: Training and operating industrial robots. https://patents.google.com/patent/US8965580B2/ , last accessed June 2020, 2015.
65. Brooks, R.A. Machine Learning Explained. https://rodneybrooks.com/forai-machine-learning-explained/ , last accessed June 2020, 2017.
66. Brooks, R. Avoid the Cerebral Blind Alley. Response to: Is the Brain a good model for machine intelligence? *Nature* **2012**, *482*, 462.
67. Pearl, J. The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM* **2019**, *62*, 54–60.
68. Brooks, R. The Seven Deadly Sins of AI Predictions. Mistaken extrapolations, limited imagination, and other common mistakes that distract us from thinking more productively about the future. *MIT Technology Reviews. Online https://www. technologyreview. com/s/609048/the-seven-deadly-sins-of-ai-predictions/* **2017**, *6*.
69. Strannegård, C.; Svangård, N.; Lindström, D.; Bach, J.; Steunebrink, B. The animat path to artificial general intelligence. Workshop on Architectures for Generality and Autonomy, IJCAI-17, 2017.
70. Wiedermann, J.; van Leeuwen, J. Understanding and Controlling Artificial General Intelligent Systems. 10th AISB Symposium on Computing and Philosophy, in AISB Symposium X; , 2017.
71. Sloman, A. A Philosopher-Scientist's View of AI. *Journal of Artificial General Intelligence* **2020**, *11*, 91–96.
72. Brooks, R.A. My dated predictions. https://rodneybrooks.com/my-dated-predictions/ , last accessed June 2020, 2018.
73. Russell, S.; Norvig, P. *Artificial intelligence: a modern approach*, 3rd ed.; Prentice Hall, 2009.
74. McCulloch, W.S., Ed. *Embodiments of Mind*; MIT Press: Cambridge, MA, 1965.