

An Integrated Understanding of AI

Draft 20 November 2023

Andrew Basden

I was working in AI (artificial intelligence) in the early 1980s but it was a different AI - and yet the fundamental issues are the same now as then. I want to share with you a way of understanding AI that is not well-known but I have found very helpful and fruitful [FOOTNOTE: Understanding AI].

Though interest in AI today is at fever pitch, especially among politicians, business people and the media, and among academics who want to get money from them, the debate is fragmented and often based on prejudice, spectacle and misunderstandings. The way of understanding AI that I have found offers an integrative, holistic picture of both technology and use together, and is based on understanding the nature of reality itself. It is a philosophical way of understanding that happens also to be intuitive. It emerges from the philosophy of Dooyeweerd, a mid-twentieth century Dutch thinker.

For decades two main questions were asked about AI,

- ◆ Q1: “Could computers ever become like humans?”
- ◆ Q2: “Will AI take over the world, making humans extinct or allowing us to live without working?”

Elon Musk recently claimed that AI will do all our jobs; that claim was also made in the 1970s too! Since the automated cars and ChatGPT burst on the scene, however, other questions have entered the debate, such as:

- ◆ Q3: “Will AI (ChatGPT) write essays for students?”
- ◆ Q4: “Will automatic cars kill cyclists who are pushing their cycles?” (one did)
- ◆ Q5: “Surely AI is better than us at detecting cancers in x-rays / finding new chemicals / etc.”
- ◆ Q6: “Will AI recognise my face and put me at risk?”
- ◆ Q7: “Will AI change society?”

The questions express different issues. Q1 and Q2 concern what is called general AI, one overtly philosophical, one about distant possibilities. Q3 to Q7 are more prosaic questions about particular applications. About applications, Q3 to Q7 differ from each other, with Q3 being about capability, Q4 about AI going wrong, Q5 about which applications might be possible, Q6 about how we use AI to benefit or harm, and Q7 similar but at a society level. Q4, Q6 and Q7 have an important normative thrust.

There seems to have been no conceptual framework that can inform our thinking about all six questions together. When we address such questions, we assume a conceptual framework to help us do so, which itself is based on a set of philosophical ideas. To date, different philosophical ideas inform the debate about the different questions, so that, in the main, the questions are addressed in isolation from each other. I have discovered that Dooyeweerd’s philosophy [Dooyeweerd 1955], which comes from a different root than most, can allow us to address them all together.

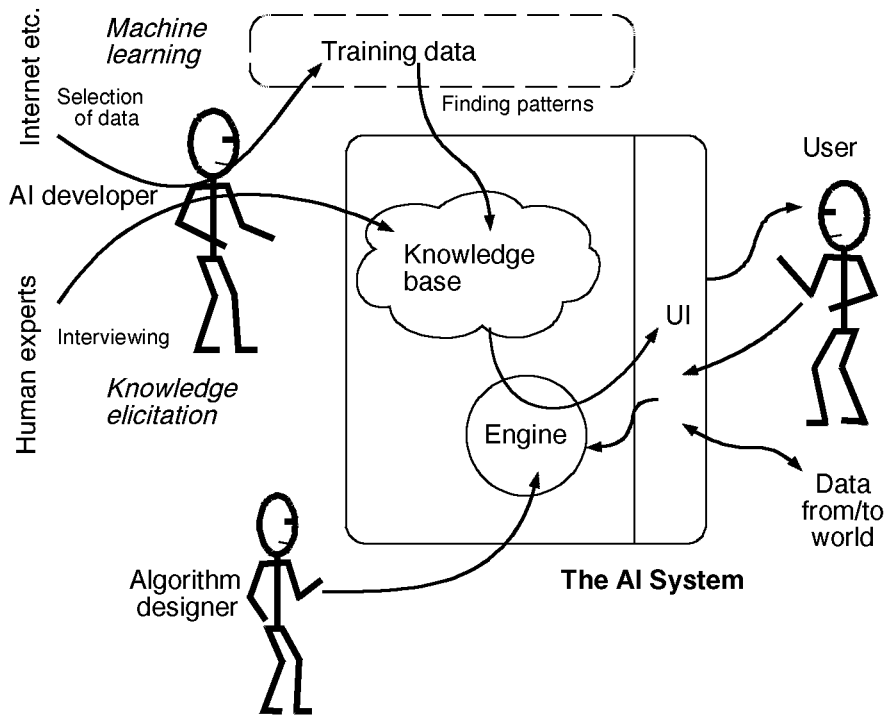
Often, at this point I would open a section explaining Dooyeweerd’s philosophy, but that is not necessary because much of his philosophy is intuitive if we maintain an open mind. Instead I will introduce it bit by bit as we need it for addressing the different questions. This article is in three sections, all brief but hopefully complete enough:

- ◆ Part 1 explains how AI works - both kinds of AI, to give us perspective.
- ◆ Part 2 addresses the application questions, Q2 to Q7.
- ◆ Part 3 addresses the abstract, philosophical question, Q1, after we have understood the realities of AI.

This article is intended to be helpful to the intelligent reader who knows little about AI but wants to understand more. I admit having a Christian perspective, but it is not the normal one, and should be interesting to most people.

Part 1. How AI Works

The following figure shows roughly how AI works. The AI system is a software engine operating with a knowledge base, interacting with users via a user interface (UI) and sometimes data from the world via sensors or databases. (In automated AI the UI might be only a start/stop button, a few controls and data from sensors, but in most AI, like ChatGPT, there is more 'dialogue' between users and AI systems.)



The knowledge base encapsulates knowledge about how it should operate in its intended application, based on various technologies, like inference nets, sets of logical statements, sets of associations, or so-called neural networks. The engine is designed and written to process the encapsulated knowledge according to the technology employed so as to respond to users (or the world).

For example, at the core of ChatGPT is a huge matrix of probabilistic associations between phrases and words found in billions of statements taken off the Internet (with a lot more around this, such as images). Its engine uses this both to understand user questions or instructions and to generate replies or even essays [FOOTNOTE: ChatGPT].

Two Kinds of AI

There are two kinds of AI, two ways in which the knowledge base can be

constructed, in which the AI developer operates in a different way: human knowledge elicitation and machine learning.

In my early work as AI developer in the 1980s, we would manually build the knowledge base by interviewing human experts and expressing the elicited knowledge in an appropriate computer language. Knowledge engineering, as it was called, was a labour-intensive process, in which good knowledge engineers would winkle out tacit knowledge and rare exceptions and incorporate them into the knowledge base.

Today's machine learning AI (MLAI) bypasses the human processes of eliciting and expressing knowledge, by detecting patterns in masses of training data supplied to it by AI developers, such as from Reddit in the case of ChatGPT. [FOOTNOTE: MLAI] I like the explanation given by Paul McCartney [Kraftman 2023] of how they used MLAI to extract John Lennon's voice from a poor quality recording; they told the AI system,

"That's voice. That's guitar. In this recording, lose the guitar."

Why Humans Are Important

How well AI works depends on the quality of knowledge in its knowledge and, of course, on the engine processing this correctly. Since human beings design both engine (algorithm designer) and knowledge base (AI developer), and also use the AI system, even if indirectly, AI cannot be properly understood without taking human intention and interpretation into account.

The quality of early AI depended on sensitive elicitation and close relationships of trust with experts. Sadly, because AI became fashionable, many became knowledge engineers who would be less careful, so that many AI systems did not work well. Quality of MLAI depends on careful selection of training data and of parameters by which to learn patterns.

In both kinds, the quality of the knowledge base is a human responsibility. Part 2 helps us understand what that quality is, and more about the ways humans are important.

Part 2. Understanding What AI Can And Should (Not) Do

AI (Artificial Intelligence) can beat us at Go and Chess. AI let an automated car kill a cyclist. AI can analyse X-ray screens very well. ChatGPT can write essays for students, but they are bland and full of errors ("hallucinations").

As we pointed out above,

- ◆ Q2: "Will AI take over the world, making humans extinct or allowing us to live without working?" is about the ultimate possibilities inherent in AI, and whether AI will ever be able to take over from humans.
- ◆ Q3: "Will AI (ChatGPT) write essays for students?" calls us to understand the capability of AI.
- ◆ Q4: "Will automatic cars kill cyclists who are pushing their cycles?" calls us to understand AI going wrong.
- ◆ Q5: "Surely AI is better than us at detecting cancers in x-rays / finding new chemicals / etc." calls us to understand in which kinds of application AI can be successful
- ◆ Q6: "Will AI recognise my face and put me at risk?" calls us to understand how we use AI to gain benefit - or harm.
- ◆ Q7: "Will AI change society?" calls us consider impact of widespread use of AI on society and planet, and structures of society.

Q2 can only be addressed after addressing Q3 to Q7.

Q3. What Makes AI Capable?

The capability of an AI system comes mainly from its knowledge base encapsulating laws and information that are meaningful in aspect(s) of reality relevant to its application: spatial aspect for Chess AI, kinematic aspect for automated cars and lingual aspect for ChatGPT, for example.

But what aspects are there? Dooyeweerd carefully delineated fifteen that seem to be irreducible to each other (i.e. cannot be explained in terms of each other nor inferred from each other, just as, in architecture, the east and south aspects of a building cannot be inferred from each other). The following table lists his aspects, along with what the laws of each are about and some typical AI applications that are mentioned here, in which the aspect is central. [FOOTNOTE: Laws]

Aspect	Laws to do with ...	Applications mentioned
Quantitative	Arithmetic	-
Spatial	Spatiality	Chess; X-ray analysis
Kinematic	Movement	Automated cars
Physical	Energy, forces, causes Physics, chemistry	Molecule design
Organic / Biotic	Life functions, cells, organs, Organisms, Ecosystems	Face recognition
Psychic / Sensitive	Stimulus-response	Voice extraction
Analytic	Concepts, analysis, logic	Business analysis
Formative	Design, planning, forming	-
Lingual	Signifying, speaking, writing	ChatGPT
Social	Associating, institutions	-
Economic	Resources, frugality	(Business analysis)
Aesthetic	Harmony, enjoyment	-
Juridical	Rightness Reward, punishment	-
Ethical / Attitudinal	Self-giving love	-
Pistic / Faith	Belief, commitment Ultimate meaning	-

'Human' aspects

'Successes'

So, for Chess, AI must have a good 'knowledge' of the laws of the spatial aspect and ChatGPT, of the lingual aspect, so that it can function in them.

However, it is more complicated than that because Chess AI must have *some* 'knowledge' that is meaningful in other aspects, such as of movement (kinematic aspect) human goals and strategy (the formative aspect). ChatGPT must have *some* 'knowledge' of the formative aspect (structure of language), analytical aspect (distinguishing words, phrases and part-words from each other: vocabulary etc.), psychical aspect (especially for colour in pictures), spatial aspect (in pictures), social aspect (it has a database of people and their relationships), and a few others. Such we will call the secondary aspects, because they are there to support its operation in its main aspect. Some AI systems might have more than one main aspect; we need not be dogmatic about which are primary and secondary; the idea of primary aspect is here to help us understand, especially for Q5.

How can ChatGPT write essays, for example? ChatGPT analyses user's instructions or questions, and generates the text of the essay. Both operate according to the laws of the lingual aspect, which are encapsulated as host of probabilistic degrees measuring how much each word is meaningful in more than 12,000 ways. With this, ChatGPT's algorithm is designed to perform conceptually simple mathematical matrix operations by which the relationships among words can be reasoned about, for example which words tend to follow which in various contexts and which words are synonyms for each other.

[FOOTNOTE: How ChatGPT works].

It is not divulged what those 12,000 ways are but we may expect each to represent a different permutation of the fifteen aspects. To Dooyeweerd, aspects are "modalities of meaning" that are irreducible to each other and yet intertwine with each other. In actuality ('real life') all aspects operate together; they are aspects of our activity and being, not activities and beings themselves - ways in which our being and functioning can be meaningful and which combine to define the meaning of a thing.

This massive knowledge base was constructed by ChatGPT reading vast amounts of Internet content (175 billion pieces as of November 2023). Since all these pieces are results of humans functioning in the lingual aspect (consciously or subconsciously), they together express human beings' functioning in the lingual aspect. In 1980s AI, the laws of the lingual aspect would have to be elicited and encapsulated in the knowledge base explicitly and manually.

But why does AI make mistakes, such as in automated cars not recognising a cyclist pushing a bicycle. or ChatGPT offering its famous "hallucinations"? That is the issue addressed in Q4.

Q4. Why Does AI Go Wrong?

There are several reasons AI goes wrong. One is errors in user input or world data. Three others arise from deficiencies in the encapsulated knowledge.

1. Erroneous knowledge in the knowledge base. Because human writings from the Internet contain errors, ChatGPT 'learned' erroneous patterns that generate "hallucinations". Also, since its word associations are probabilistic, it sometimes selects inappropriate ones.
2. Missing knowledge: minor biases. Tacit knowledge and rare exceptions are often absent from a knowledge base. In knowledge elicitation, a good analyst will deliberately seek these out but MLAI learns patterns statistically. There is often not enough training data to learn rare patterns reliably, such as cyclists pushing rather than riding bicycles.
3. Missing aspects: major biases. Omitting a whole aspect omits a whole swathe of knowledge that is meaningful in that aspect. Whole aspects might be missing if the AI developer fails to recognise their relevance and so does seek them or provide training data about them. This becomes problematic when AI is used in different contexts. Most training data for ChatGPT was written by affluent people in the Global North, in which some aspects important elsewhere have been undervalued.

It is the AI developer who is responsible for ensuring high quality knowledge bases. This becomes more challenging in later-aspect applications, as addressed in Q5.

Q5. In Which Applications Can AI Work Well?

In which applications AI is likely to work well (now and in future), can be understood via aspects. The laws of earlier aspects are easier to encapsulate in a knowledge base reliably. for two main reasons. One is that the laws of earlier aspects are more determinative so that, for example, $3 + 4$ is always 7 (law of quantitative aspect), whereas a question might be answered in several different ways (lingual aspect).

The other is that the laws of earlier aspects act as a foundation for those of later aspects, so, in principle, encapsulating knowledge of later aspects requires us to encapsulate laws of all earlier aspects too. Laws of physics depend on three earlier aspects, those of lingual, on eight. Moreover, the middle aspects of human individual functioning are influenced by later aspects too, which can also need encapsulating (e.g. ChatGPT's social database).

Therefore AI tends to work more reliably, and have more successes, in applications governed by the earlier aspects, than those governed by later aspects (see Table 1). X-Ray analysis (spatial aspect) is more reliable than is ChatGPT (lingual). Those who extrapolate from current successes in AI to "AI will soon be able to do everything" fundamentally misunderstand AI.

However, full reliability is not always needed where AI *assists* rather than *replaces* humans - the next question, Q6.

Q6. How Do We Use AI for Benefit Not Harm?

Whether AI face recognition is beneficial or harmful depends, not just on the AI working properly or wrongly, but the role it plays and whether it is used with evil intent, carelessness or good intent. Nor will AI do all our jobs, as Elon Musk believes; similar predictions were made in the late 1970s!

Roles: Most popular discussion presupposes AI replacing humans, but AI can also assist humans. During the 1980s, I was involved in an AI system to advise managers on the strength of business sectors - analytical and economic aspect application. From information supplied by managers, it estimated sector strengths but then actively encouraged them to disbelieve it rather than accept its answers, inviting them to explore differences between their and its views. This revealed things they had overlooked, thus refining their knowledge. Knowledge refinement is the very opposite of AI replacing humans [FOOTNOTE: Roles of AI].

Intent, at two levels: Whatever role, is AI used with good intent, evil intent or carelessness? Are decisions to invest in or deploy AI made with responsibility and wisdom, or with self-interest and fear of missing out?

Q7. How Might AI Affect Society and Planet?

As I argued for information technology in general in chapter 5 of Basden [2018], there are two societal issues. The first is when use in an application becomes widespread: its impact, for good or ill, is multiplied by millions. For example billions of miles being driven each year contributes one third of the climate change emissions we face and also, researchers tell us, makes society more individualistic and selfish. The widespread use of AI to choose advertisements for us on social media: what impact does that have? We might not be able to predict this in the usual ways, but Dooyeweerd's aspects at least enable us to separate out kinds of impact - such as biodiversity, health, mental stress, friendliness, waste of resources, trust, etc. - and ensure that none are overlooked.

The second societal issue is societal structures that constrain and enable how we live. The most visible of these is legislation, which defines some things as legal and others as illegal (a distinction that is meaningful in the juridical

aspect). But there are two other kinds of societal structure, meaningful in the ethical and pistic aspects: attitude that pervades a society (self-giving and open, or self-protective and selfish) and mindset that prevails throughout a society (what society deems most meaningful, to be aspired to, to be expected or taken for granted, to be sacrificed for, etc.). Will AI affect these three kinds of structure, and thus our lifestyles, and thus our impact on planet? Most discussion of societal structure focuses on the juridical, of rights and responsibilities, and the tool to address this is legislation shaped to encourage or curb AI. Legislation is obvious, but attitude and mindset are hidden and less often discussed, and yet arguably more powerful in their effects.

Q2. Will AI Take Over From Humans?

No. Because, to do so, it would have to (a) have encapsulated in its knowledge base the laws of every aspect (b) have done so completely and without errors or biases, including cultural. For the reasons discussed above I do not believe this is possible.

The danger from AI, in my opinion, is not AI capabilities but human sin. Humanity will tend to use AI in ways that are “affluent, arrogant and unconcerned”, which is the reason Sodom was destroyed and Judah was exiled [Ezekiel 16:49]. This attitude and mindset can affect all three human activities around the AI system, algorithm design, AI development, and AI use and deployment. [FOOTNOTE: Sam Altman] Issues of climate change, biodiversity and the Global South are largely overlooked so far but, I submit, are more important issues in God’s eyes, and for our future, than AI capability.

On the Ethics of AI

Increasing numbers of people are discussing what is called the ethics of AI. Sadly, this discussion often takes place in a different room from the technology and capabilities of AI, but in Dooyeweerd’s view that we adopt they cannot be so separated. Capability and technology and ethics are inescapably intertwined. This is because to Dooyeweerd the fifteen aspects all define various kinds of Good and, from the biotic aspect onwards, a corresponding Evil or Harm, and from the analytical onwards, authoritative normative guidance for human living.

In asking Q4, “Why does AI go wrong?” we are already presupposing some notion of wrongness, but usually implicitly. Dooyeweerd enables us to make this explicit. AI developers and algorithm designers are not “rational economic actors” but fully human beings, and function in all aspects in doing their development and design. Do they function transparently or deceitfully (lingual aspect)? Cooperatively or competitively (social aspect)? justly or unjustly (juridical aspect)? Generously or selfishly (ethical aspect)? And so on. All these affect the quality of their knowledge bases and algorithms, often in subtle ways that only become evident later on.

Q6, “How Do We Use AI for Benefit Not Harm?” is about the similar multi-aspectual functioning of the users and deployers of AI. It is obvious that using face-detection AI to find someone you want to kill is evil, but why? Because of the juridical norm of justice and due. Using face-detection AI to find people who are starving so as to be able to bring food to them is good, but why? Because of the ethical norm of self-giving love. Using ChatGPT to find information to help a student write a better essay is good in the lingual aspect; using it to cheat is wrong, in the juridical aspect.

Of course ethics of AI extends to the societal level too, especially when AI use for particular applications becomes widespread. For example, might use of

AI to allow businesses to compete for selfish advantage and destroy each other be at least dubious, if not entirely evil, in the ethical aspect?

Dooyeweerd's aspects offers us a basis for thinking about and discussing such issues with greater clarity and deeper understanding.

Part 3. Can AI Be Human?

So far we have blithely talked about computers doing things that are normally attributed to humans, such as analysing, understanding, composing, etc. But is it valid to do so? Are we not presupposing a fundamental similarity between AI and humanity? Of course, AI not yet like humans, but many argue that, given time, it will become so. Whether or not AI can ever become like humans is a philosophical question, Q1, which we address here.

Debate about Q1, "AI = Human?", has been debated for 70 years. It was raging in the 1980s, when I was an AI developer, but it remains unresolved. Why? Dooyeweerd, I believe, offers a route to resolving it because he investigated the very roots of philosophical debate over the past 3000 years, mainly in Western thought but also with awareness of other thought. The following is a summary of a more detailed discussion in Basden [2008, 207-220].

Why Has the Debate Never Been Resolved?

The ways the question has been posed are determined by what is presupposed most deeply meaningful in reality. These are "ground-motives", which propel a society's thinking and beliefs over centuries. Four have driven Western thought for 2500 years, of which three are dualistic, in which two poles, X and Y, are absolutely opposed, and hence can never truly be brought together. The fourth, the Biblical ground-motive, is non-dualistic (in fact pluralistic). They yield different versions of the AI question (their proponents in brackets):

- ◆ The Greek ground-motive of Mind-matter: "Computers are matter, humans are partly mind; can matter generate mind? e.g. Could a dump of my mind into Cyberspace be the real me? Could I live forever that way? [John Perry Barlow]
- ◆ The Scholastic ground-motive of Nature-supernature: "Computers are natural; humans are partly supernatural; can computers gain such supernatural characteristics?" e.g. Is the biological causality by which humans operate a kind of spark that computers can never have? [John Searle]
- ◆ The Humanistic ground-motive of Nature-freedom: "Computers are machines, determined; humans are partly free; can freedom arise from determined causality?" e.g. Could Emergence Theory explain it [Allen Newell, Systems theory] Or Quantum mechanics?

Posing the AI question in any of the dualistic ways is ultimately fruitless because each presupposes the very opposition that it tries to overcome. Dooyeweerd argued that the three dualistic ground-motives have misled philosophy and science into many dead-ends, just as is happening with the AI question.

- ◆ The Biblical ground-motive of Creation-Fall-Redemption recognises, and encourages us to explore, multiple ways in which the Creation - involving human, animals, plants and inanimate things (including machines such as computers) - is Meaningful and Good (works well).

The dualistic ground-motives force debate into opposition; the Biblical ground-motive gives a basis for integration of views. It was this that motivated Dooyeweerd and his colleague Vollenhoven to investigate the diversity of

meaningfulness rather than either trying to reduce it to one aspect [FOOTNOTE: Reductionism] or presupposing human subjectivity as the sole source of meaningfulness, hence arbitrary.

The Chinese Room Thought Experiment

In 1990 John Searle suggested the **Chinese Room thought experiment**, to demonstrate, he thought, that AI can never be human, but there is a flaw in his argument that most people, even his opponents, did not see. Strictly, he sought to demonstrate that claims that appropriately programmed computers can genuinely understand (possess intentionality) are baseless. His argument is as follows:

To summarise: Suppose I do not understand Chinese, and cannot even recognise Chinese writing from any other shapes. I am in a room with a batch of Chinese writing. From time to time more pieces arrive through a hole in the wall. I also have a rule book in English (which I understand well) that tells me how to reply (by drawing shapes) to each received pattern on the basis of formal properties like its shape and which take into account all previous patterns received and sent. Also, occasionally, I receive questions in English, and reply to those. "Where," asks Searle rhetorically, "in this room is the understanding of Chinese? And how does it differ from my understanding of English?" He argued that computers running a program are like following the rule book, and cannot understand in the way human beings do.

Searle argues that biological causality is necessary for understanding, and that the physical causality of computers can never achieve this; humans operate by one while computers operate by the other. In effect, physical causality is 'lower', 'natural' while biological causality is 'higher', 'supernatural' from the perspective of the natural; the second ground-motive above. However, there is also another philosophical presupposition, which infects all three dualistic ground-motives.

Various counter-arguments have been attempted by AI supporters, of which six kinds are found in Boden's [1990]:

- ◆ The systems reply, that the property of understanding Chinese is an emergent property of the system of room, rules, me, etc.;
- ◆ The robot reply, that understanding involves action in the world (in which the symbols I send out go to robotic arms etc.);
- ◆ The brain stimulator reply, that all we need is for the program to simulate the operation of brain cells rather than rules directly (which is the principle on which machine learning neural nets is based);
- ◆ The combination reply, that putting all of these together is enough for the AI to have the property of genuine understanding;
- ◆ The other minds reply, that we cannot know what is in another mind except by the behaviour we see, so if the Chinese Room behaves aright we may say it understands Chinese;
- ◆ The many mansions reply, that eventually we will build computers with the right type of causality and these will truly understand.

Boden records Searle's reply [p.72ff], countering all of them successfully. The debate continues. [FOOTNOTE: Chinese Room]

Notice how Searle's challenge is cast in terms of an entity with some innate ability or property of "understanding", and all the attempted counter-arguments are cast in the same terms. There are fundamental problems with doing this. A philosophical one, introduced to the AI community early on [Hirst 1991] is that

existence is not as straightforward as we assume (e.g. in what sense does Gandalf exist, or a square circle not exist?). A more practical problem is that any attempt to ascribe a property X to AI, such as understanding or intelligence, is countered by "But that is not real X!" End of discussion! Discussion became sterile. The root of the problem is what Dooyeweerd called the Immanence Standpoint [FOOTNOTE: Immanence Standpoint], in which we presuppose the self-dependence of entity: it 'just is', without reference to anything that transcends it; this, he argued, has misled philosophical thinking for three millennia.

Instead of existence and properties, Dooyeweerd grounded his philosophy on meaning, on which existence and properties depend, and which in turn depends on an Origin of Meaning, i.e. a Creator. It was this that motivated him to explore the diversity of meaningfulness that we encounter, coming up with his suite of aspects.

With this, we can suggest an answer to Searle's challenge and also a way to resolve the fruitless dualism.

Addressing "AI = Human?" with Dooyeweerd

There is one obvious answer to Searle's "Where is the understanding of Chinese?" which all miss: in the book of rules. Why did they miss it? Because of their focus on properties of entities; the content of the rules is meaning, and meaning is not strictly a property.

And notice how this implies human beings: was it not human beings that wrote the book of rules? The book of rules is the Chinese Room's knowledge base in the earlier figure.

This shift to casting the question in terms of meaningfulness rather than being (with innate properties or possibilities), opens up a different way to address Q1, whether computers can be like humans. It recasts the question as "Is it meaningful to say that computers, like humans, function in aspect X?" When we do this, we find two ways of answering the question:

- (a) an 'everyday' way in which computers and humans operate together as part of Creation (the humans including designers, fabricators, programmers and users), and
- (b) a narrower, theoretical way, in which we take humans completely out of the picture. We treat the computer as a mass of silicon, various doping elements, copper, plastic, etc., all arranged in certain spatial arrangements and subjected to certain electromagnetic forces. (The reason why they are arranged this way is, in this view, irrelevant.)

[Footnote: In philosophical terminology used by Dooyeweerd, (b) is subject-functioning and (a) is any meaningful functioning whether as subject and/or object.]

In the first four aspects, answers to both (a) and (b) are "Yes" for both computers and humans. For example computers and humans consume energy (and thus emit greenhouse gases), occupy space, and so on. In these four aspects, computers are like humans. In subsequent aspects, however, the answer is "Yes" if we take humans into account (version (a)), and "No" if we do not (version (b)). See the right-hand columns in the following table.

Aspect	Kernel	(With humans) computers can ...	(Without humans) computers can ...
Quantitative	Amount	calculate	have amount of memory
Spatial	Extension	(find cancers in) x-ray plates	take up space
Kinematic	Movement	(find) routes; (play) Chess	work dynamically
Physical	Energy	(design) new chemicals	consume energy; be heavy
Biotic/Organic	Life	self-repair; (diagnose) illness	
Psychic/Sensitive	Feeling	sense and respond	
Analytic	Distinction	think and analyse	
Formative	Shaping	plan; manufacture	
Lingual	Signification	converse	
Social	'We'	bring people together	
Economic	Frugality	save money	
Aesthetic	Harmony	create music, paintings	
Juridical	Due	help find legal loopholes	
Ethical	Self-giving	help care for someone	
Pistic	Faith	assume something	
(Brackets indicate an additional aspect, e.g. analyse.)			

The answer is “Yes” in (a) because we assign meaning from later aspects to the physical operation of the computer: the way the electromagnetic fields vary and to their spatial arrangements. It is the fabricators’ intention to build a computer, which is the reason why the various chemical elements are arranged spatially they way they are. It is the designers’ and programmers’ intention to produce an application, such as ChatGPT, that is the reason for the initial arrangements (at switch-on and application loaded) of electromagnetic forces (in what fabricators would call the computer memory). It is the users entering text into ChatGPT that is are the reason for how those forces vary through time. See column 3.

The answer is “No” in (b) because, in that view, the aspects that make intention to build a computer, write ChatGPT and seeking answers, meaningful are irrelevant and what happens is described purely in terms of physical forces and energy, and their spatial arrangement and movement [FOOTNOTE: Bits]. See column 4.

Concluding Remarks

We have addressed a broader range of questions about AI than is usual, and found that Dooyeweerd’s philosophy is able to help us address them all. This approach therefore offers an holistic way of understanding AI that exhibits an innate harmony, and it philosophically sound. It brings together the two types of AI, the technical with the ethical issues, impact on individuals and society, and many different kind of applications - ChatGPT, x-ray analysis, automated cars, Chess, and so on. And it does so in ways that respect their differences.

At the core of this approach is Dooyeweerd’s suite of aspects, which is the conceptual tool we have employed. It has proven successful in research and practice in many areas [Basden 2020, especially Chapter 11]. More than that, Dooyeweerd was clear that the kernel meanings of aspects are better grasped by

intuition than in a theoretical attitude of thought - which implies that it need not take philosophical experts to grasp this explanation of AI.

Footnotes

Note on Understanding AI. The understanding of AI presented in this article is an amalgam of two sources, one being three blogs published by Faith in Scholarship, the other being my book [Basden 2018] on *Foundations of Information Systems: Research and Practice*, in which I worked out an integrated, holistic understanding of information technology and digital systems, of which of course AI is a species. If you want to take the ideas in this article forward, please read that book. Please contact me if you have problems in doing so.

Note on ChatGPT and How It Works. For an excellent, accessible explanation of how ChatGPT see [Lee & Trott 2023].

Note about MLAI. The knowledge base in machine learning AI (MLAI) is usually based on neural net technology or associations.

Note About Dooyeweerd's Aspects. Dooyeweerd's fifteen aspects may be explored by going to the aspect 'home page' at "<http://dooy.info/aspects.html>" and a summary at "<http://dooy.info/aspects.smy.html>". The fifteen aspects are Dooyeweerd's best guess at the complete range of ways in which things may be meaningful. Other suites of aspects could be used, but Dooyeweerd's is most complete and most philosophically sound; see "<http://dooy.info/compare.asp.html>". Dooyeweerd was clear that no suite of aspects, including his own, can ever be treated as a final truth, so we take them on trust as a conceptual tool to help us think.

Note about Laws. Laws here are not like laws of a land nor social norms, but laws that govern how things function. The law of gravity, for example, is a law of the physical aspect, and it enables masses to stay together. The lingual aspect has laws that enable language to occur, which are deeper than any one language group. Laws of the later aspects are non-determinative, but they guide towards what is Good.

Note about Roles of AI in Use. Basden [1983] outlines eight roles in which AI could be used and be beneficial. Strangely, there has been little discussion of roles since then, but most of the roles still apply today.

Note on Sam Altman. When I first investigated ChatGPT and how OpenAI was changed by its founder and CEO, Sam Altman, to be non-open, I suspected something of this attitude of "affluent, arrogant, unconcerned". It was therefore no surprise to me to hear today (20 November 2023) that he was fired for not being entirely honest in his dealings with the OpenAI Board.

Note on Reductionism. Reductionism has several forms, discussed in Clouser [2005], including treating only one thing or aspect as valuable or meaningful, such as reducing everything to money, and trying to explain the entire complexity we encounter in terms of one aspect, such as materialism and evolutionism do. Trying to break out of reductionism is system thinking, which tries to accept multiple aspects. Dooyeweerd offers a useful conceptual tool to help this.

Note on Chinese Room. For a fuller discussion, see Basden [2008, 210-216].

Note on Immanence Standpoint. The Immanence Standpoint, as Dooyeweerd called it, a presupposition as to the deepest idea of what reality is like. The ancient Greeks presupposed "It exists" to be the most fundamental thing we can say about something, and existence was presupposed to be self-explanatory and self-dependent. But as Hirst [1991] points out existence is neither. Clouser [2005] offers a good explanation of this, especially the idea of self-dependence.

Dooyeweerd rejected the Immanence Standpoint, holding that existence always presupposes meaning, To say that a poem exists is to say that something is functioning in ways meaningful in the aesthetic aspect (and others).

Note on Bits. It is commonly thought that “the computer is only ones and zeros” (which are called “bits” in digital systems). This is not strictly true. a bit of value 1 can be implemented electronically as a voltage of 3v or 5v or 12 v, as a current flowing, or as a phase change in an AC current. The bit-value is an attribution to a physical phenomenon from the perspective of the psychical aspect, in which signals are meaningful. Moreover, there are also analog computers which do not operate with bits. To speak about bits is to describe the computer from the perspective of the psychical aspect.

References

Basden A, (1983). On the application of Expert Systems. *Int. J. Man-Machine Studies*, 19:461-477. “<http://kgsvr.net/andrew/-p/ai/Basden83-ApplicES.pdf>”

Basden A. 2008. *Philosophical Frameworks for Understanding Information Systems*. IGI Global Hershey, PA, USA.

Basden A. 2018. *Foundations of Information Systems: Research and Practice*. Routledge, London, UK.

Basden A. 2020. *Foundations and Practice of Research : Adventures with Dooyeweerd’s Philosophy* Routledge, London, UK.

Boden MA. 1990. *The Philosophy Of Artificial Intelligence*. Oxford University Press, UK.

Clouser R. 2005. *The Myth of Religious Neutrality; An Essay on the Hidden Role of Religious Belief in Theories*. University of Notre Dame Press, Notre Dame, Indiana, USA.

Dooyeweerd H. 1955. *A New Critique of Theoretical Thought*, Vol. I-IV, Paideia Press (1975 edition), Jordan Station, Ontario.

Hirst G, (1991), “Existence assumptions in knowledge representation”, *Artificial Intelligence*, **49**:199-242.

Kraftman T. 2023. *Paul McCartney is using AI to create a “final Beatles record*. Available at <[a href = “https://guitar.com/news/music-news/paul-mccartney-is-using-ai-to-create-a-final-beatles-record/”](https://guitar.com/news/music-news/paul-mccartney-is-using-ai-to-create-a-final-beatles-record/)> Guitar.com.

Lee TB, Trott S. 2023. A jargon-free explanation of how AI large language models work. Available on <[a href = “https://arstechnica.com/science/2023/07/a-jargon-free-explanation-of-how-ai-large-language-models-work/.html”](https://arstechnica.com/science/2023/07/a-jargon-free-explanation-of-how-ai-large-language-models-work/.html)> Ars Technica.

Copyright (c) Andrew Basden 2023.

Created: 20 November 2023.

Last updated:

</body> </html>