The Identity of Persons and Turing Machines

1 Introduction

A fundamental question in computer research is whether a person is a computable entity, i.e. whether the time dependent relation which connects various states of a person (personal identity) is in principle Turing computable.

Seeking to reject this idea, I shall in §2 characterise three notions: personal identity, deductive Turing machines, and information theory. In §4 I shall detail how personal identity can be seen as information transmission and thereby serve as the input to a Turing machine. In §5 I shall interpret non-branching accounts of personal identity in light of information transmission and show them decidable. In §6 I shall argue that non-branching accounts are inferior to branching accounts due to transitivity issues. In §7 I shall interpret branching accounts of personal identity in light of information transmission and show them undecidable. I shall thus conclude that if the identity of persons is branching, then it cannot be decidedly computed by deductive Turing machines.

2 Background

2.1 Motivation

Present endeavours in artificial intelligence are progressing towards approximating human intelligence. 'Classical logic architectures' have been superseded by connectionist networks (neural models constituted by multiple connected and weighted units) that are challenging human performance in areas like vision[1], language production[2], and domain specific judgement[3]. Regardless, AI algorithms are built in consistent axiomatic systems which are subject to Turing computability. Whereas narrow AI has been the focus of contemporary research, the ultimate goal for artificial agents lies in artificial general intelligence, i.e. fully modeling the human persona by artificial means. This endeavour evokes the question whether the person itself is in principle a Turing computable object or whether it holds a more privileged metaphysical status.[4] Whilst initially a perplexing thought, the person as computable has an established history in

[1]Thrun 2017
[2]Brown 2020
[3]Schrittwieser 2020
[4]B. C. Smith 2019, 71

cognitive science where both connectionist[5] and dynamical systems[6] approaches have been deployed toward a computational theory of mind. More practically, development in the intersection between computer science and neuroscience has stood to pioneer scientific endeavours to support these ends.[7] By this I shall investigate a novel intersection between the *identity* of persons and theoretical computer science to answer whether personal identity is in principle Turing computable, i.e. a computable relation waiting to be quantised and replicated by artificial means.

Before answering this question, I shall provide an analysis of personal identity in terms of information theory to sharpen its definition for serving as the input to a Turing machine. The original question then enquires whether a Turing machine computing personal identity will terminate in finite time, i.e. whether personal identity is decidable. If positive, this would entail that the identity of persons is a computable object.

2.2 Personal identity and continuity

Take three initial conditions to constitute personal identity:

- *Identity*: $x = y$ iff, $\forall P(Px \leftrightarrow Py)$.
- *Personhood*: the persistence properties P under which a person x at t and $t + 1$ is the same person.
- *Personal identity*: $x_t = x_{t+1}$ iff, $\forall Px_t \forall Px_{t+1}(Px_t \leftrightarrow Px_{t+1})$.

From this I take personal identity as a map of properties projected between one state of a person, x_t, to another, x_{t+1}, $m : Px_t \to Px_{t+1}$.

There are broadly two routes to describe what these necessary and sufficient properties of persons are:

1. *Psychology*: Ps are psychological properties.[8]
2. *Biology*: Ps are biological properties.[9]

For the purposes of this essay I shall follow (1): personal identity as psychological continuity, since it remains the dominant view in the literature. [10]

My argument turns on a divide between psychological continuity accounts: branching and non-branching. Branching is the idea that the timeline (branch) a person exists on whilst being identical to themselves can split into multiple timelines (sub-branches) if there are disruptions to identity. Branching

[5] Ramsey 1990
[6] Gelder 1992
[7] Kriegeskorte 2018
[8] Parfit 1986
[9] Olson 1997
[10] However my arguments are not entirely biased towards it since information theory and dynamical systems have equally been used to analyse biological complexity and cellular automata, Tanevski 2016.

accounts want to honour this picture and generally stipulate a secondary relation (survival) which ensures the person persists on the new sub-branches.[11] Non-branching accounts deny such branching and instead aim to reinterpret or constrain one of the personal identity conditions.[12]

2.3 Turing machines and decidability

Let us define some primaries of Turing machines.[13] I shall assume Turing's initial definition of them as deductive Turing machines.

Starting with functions:

> *Partial Recursive Functions*: The partial recursive functions are those that can be defined from the basic functions[14] by a chain of operators.[15]

These underpin the Turing-computable functions. Turing computation is defined as follows. Start with an arbitrarily long line of 'cells' which can contain either of two values (0, 1), or else is blank. Imagine a machine capable of perceiving one cell at a time. At any cell the machine has a five membered action space:

- 0: Write a 0.
- 1: Write a 1.
- B: Write a blank.
- L: Move one step to the left.
- R: Move one step to the right.

Define an instruction-quadruple (i-quadruple) to be an ordered quadruple $< q_1, S, A, q_2 >$, whose elements are:

1. q_1: a non-zero number.
2. S: the contents of the scanned cell (0, 1, blank).
3. A: the action space (0, 1, B, l, R).
4. q_2: a number indicating the next instruction to execute.

For example, we can imagine the i-quadruple at q_1 as stating: if the scanned cell contains S, perform the action indicated by A, then go to the instructions with label q_2; if q_2 is 0, then halt the execution of the program.

By this characterisation we can establish two central tenets:

[11]Parfit 1986, 200
[12]Lewis 1976, 20
[13]For more extensive definitions see P. Smith 2013, 310
[14]Successor: $S(\mathbb{N}) = \mathbb{N} + 1$; Zero: $Z(\mathbb{N}) = 0$; Identity $I_i^k(x_1, ..., x_k) = x_i$.
[15]Composition: $f(x) = g(h(x), j(x))$; Primitive recursion: $f(x, 0) = g(x) \rightarrow f(x, Sy) = h(x, y, f(x, y))$; Minimisation: $f(x) = \mu y[g(x, y) = 0]$.

1. *Turing computability*: A Turing program is a finite consistent set of i-quadruples, and a function is Turing-computable iff there is a Turing program that computes it.

2. *Decidability*: A property/relation is decidable iff there is an algorithmic procedure that a suitable Turing machine could use to decide, in a finite number of steps, whether the property/relation applies to any appropriate given item(s).

2.4 Information

To enable the Turing computation of personal identity, we can use definitions from theories of information. Information is a general concept which denotes any amount of data in any medium. I shall define information operationally using an adapted formulation of Shannon's work.[16]

Take a communication channel between a signaler S and a receiver R. A Bit of information B passed from S to R will be disordered during transmission in proportion to the noise in the channel. If there is no noise, then B would fully resolve. If there is sufficient noise to completely block B, then no information resolves. Successful information transmission in a channel therefore abides by a probability distribution $p(B)$, where $0 \leq p(B) \leq 1$.

So, assume the *Information Definition* to be a 3-tuple (S, R, B):

- S *(signaler)*: an entity which transmits information.
- R *(receiver)*: an entity which receives information.
- B *(information)*: a piece of information measured in bits.

To compute transmission of information in a system which has a memory and evolves in time, such as a person, the formalism for dynamical systems can be used.[17] A dynamical system is a model which computes the probability distribution of a set of points evolving in time. For our purposes simply note five definitions of such systems:

- System: a set under consideration.
- Constants: members of the set.
- State ($|\psi\rangle$): a configuration of a system with the requirement that the probabilities of all sub-systems (containing constants) add up to 1.
- Density matrix (ρ_{ψ_i}): a record of the system's evolution through time which yields a probability distribution.
- Dimensionality: the number of degrees of freedom the system is defined upon.

[16] Shannon 1948
[17] Gelder 1992

Each possible configuration of the system is characterised by a state. When a computation is performed on the system, the density matrix keeps track of all possible states and yields predictions about the system's future according to a probability calculus derived from the states. The calculation of the density matrix thus determines whether the system is decidable.

3 Argument statement

In essence, personal identity is psychological continuity where psychological properties persist through time. Turing machines are deterministic algorithms which halt in finite time, i.e. are Turing computable/decidable. We seek to answer whether personal identity is in principle Turing computable. I shall argue negatively as follows:

1. Turing machines can only compute decidable systems.
2. Personal identity can be seen as an information system.
3. By (2), non-branching accounts of personal identity are decidable.
4. By (2), branching accounts of personal identity are undecidable.
5. Branching accounts of personal identity are superior to non-branching.
6. ∴ Turing machines cannot compute personal identity.

(1) is an established fact.[18] I shall argue for (2) by drawing on Benthem's account of persons as information transmission systems.[19] I shall argue for (3) by detailing how Locke's account results in a dynamical system which is decidable.[20] I shall argue for (4) by detailing how Parfit's account results in a dynamical system which is undecidable.[21] By highlighting the insufficiency of non-branching accounts contra branching ones, I shall establish (5). This will lead me to conclusion (6) that personal identity is not Turing computable.

4 Priming personal identity for computation

To bridge the gap between the informal concept of personal identity, and the formal definition of Turing machines, we need to make the definition of personal identity more precise. Towards this end I shall establish two claims:

> *Claim 1*: Psychological continuity can be seen as information transmission.

[18] P. Smith 2013, 314
[19] Benthem 2003, 10
[20] Lipton 1986
[21] Arbieto 2004

> *Claim 2*: Psychological continuity can be modeled by dynamical systems.

Let us take each claim in turn.

4.1 Claim 1

4.1.1 Persons and information transmission

Following classical interpretations of cognition as computation in cognitive science,[22] we can support Benthem's[23] characterisation of the evolution of a person's psychological properties (knowledge, beliefs, and general mental states) as information transmission.

Benthem analyses the continuous process and mutual influence of action and memory formation as information transmission. He assumes it as a three-stage process: (i) gathering, (ii) storage, and (iii) inference.

Consider an example: a person x decides to put on Chanel No. 5 rather than Miss Dior. The situation is modeled as: $C \vee D, \neg D \therefore C$. The transmission of information can be understood as: (i) x has gathered information from communication and observation about Chanel No. 5 and Miss Dior; (ii) once sufficient information has been accumulated in storage, (iii) x infers a judgement and commits to an action – in this case choosing Chanel No. 5. So, x has progressed through a cycle of information transmission.

Benthem analogises this information transmission to computation. A person's memory is an information state. When the person interacts with their environment they are exposed to information-producing events. These events transition the person from one information state to another. So, we can define information transmission as transitions between information states.

Take the example again. The information states are sets of options at any stage, i.e. {Chanel No. 5, Miss Dior}. Each option has a probability of being chosen which is determined by previous states stored in memory – in this case assume $[0.5, 0.5]$. Successive update actions on these states are then triggered by information inputs to x. This reduces the uncertainty of the options and eventually leads to an action, i.e. from $[0.5, 0.5]$ to $[1, 0]$. The outcome of this action produces new information that is stored in memory, and the cycle repeats. So, new information changes the information state, and the information state informs action which becomes new information.

4.1.2 Personal identity and information definition

We can map Benthem's characterisation onto the case for personal identity. Analysing psychological properties as information states makes the persistence of psychological properties fit the *Information Definition*:

- S (signaler): x at time t.

[22] Churchland 1989
[23] Benthem 2003, 10

- R (receiver): x at a future time $t+1$
- B (information): psychological property P, i.e. a memory.

For illustration, consider the following. A person x's memory P is preserved between two of x's information states: $Px_t \rightarrow Px_{t+1}$. The memory is of x's first date with y where x wore Chanel No. 5. If x_{t+1} remembers the meeting ($p(P) = 1$), then $x_{t+1} \neq x$. If x_{t+1} forgets the meeting ($p(P) < 1$), then $x_{t+1} \neq x$.

I shall take this characterisation to capture how information transmission can serve as a formal analysis of personal identity. So, by support of Benthem's work I determine *Claim 1* fulfilled.

4.2 Claim 2

The benefit of this information theoretic interpretation is that it allows importing formalism used to measure information transmission, i.e. dynamical systems, and deploy it towards personal identity. This in turn allows an analysis of decidability.

A dynamical system can model personal identity as follows:

- System: a person.
- Constants: all possible memories of the person.
- State ($|\psi\rangle$): one configuration of the person's memories.
- Density matrix (ρ_{ψ_i}): a record of all possible states the person's evolution can assume.
- Dimensionality: the number of possible timelines the person can evolve according to.

According to Benthem's analysis, the crucial feature of personal identity is the time-dependent conditional probability of a person forming a new memory based on past memories. This is characterised and computed by the density matrix. The dimensionality of the system scales according to the number of timelines under consideration. By these notions dynamical systems formalism can be used to model personal identity. When a system so applied is decidable, a Turing machine can compute it.

By this I determine *Claim 2* fulfilled. This establishes an information theoretic interpretation of personal identity. So analysed, personal identity can in principle serve as the input to a Turing machine. I shall now turn to analyse how differing conceptions of personal identity result in decidable and undecidable systems.

5 Computing non-branching personal identity

Non-branching accounts of personal identity result in decidable systems. I shall argue as follows:

(a) Non-branching personal identity should be modeled as a 1-dimensional dynamical system.

(b) 1-dimensional dynamical systems are decidable.

(c) $\therefore$ Non-branching personal identity is decidable.

Let us take each premise in turn.

5.1 Non-branching and 1-dimensional dynamical systems

Non-branching psychological continuity applies a 1-dimensional analysis of the persistence of temporal parts. Whilst many different forms of non-branching continuity exists,[24] we can follow Reid's interpretation[25] of Locke as a paradigm example.[26] Locke can be seen as a straightforward memory theorist: psychological continuity is a linear continuity of memory. So interpreted, every time-slice is a sequential event with a single degree of freedom: between two information states of x, either x persists as x_{t+1} or x does not, there is no alternate timeline. Psychological continuity can thereby be modeled by 1-dimensional dynamical systems.

Consider an illustrative example. At some time t, x went on a date with y wearing Chanel No. 5. At some future time $t+1$, x can be in either one of two states: (i) x remembers the date, then $x_t = x_{t+1}$; (ii) x does not remember the date, then $x_t \neq x_{t+1}$. This can be expressed as a tuple {remember, $\neg$ remember}. Assume the two states to initially have equal probability (ρ_{ψ_i}): $[0.5, 0.5]$. Successive update actions on x between t and $t+1$ then determine the result of the final outcome. For example, assume it to converge towards $[1, 0]$. In this case $x_t = x_{t+1}$.

In sum, Locke's account sees personal identity as a linear relation. Whilst each successive information state $x_t, x_{t+1}, x_{t+2}, ..., x_{t+n}$, remembers the previous states, x's personal identity is preserved. By this characterisation non-branching accounts can serve as inputs to a Turing machine. I take this to establish *premise a*.

5.2 1-dimensional dynamical systems are decidable

Moving to *premise b* we see that personal identity, as established by *premise a*, is decidable. 1-dimensional dynamical systems are decidable since 'point-to-point reachability' can be proved decidable in them. The decidabiliy problem

[24] Demarest 2015
[25] Reid 1785 (1969), III.6
[26] Locke 1694 (1975), II.27.9

was initially posed as the 'Orbit Problem' for linear sequential machines by Harrison.[27] It was subsequently shown decidable in polynomial time by Kannan and Lipton.[28] So, if personal identity is a 1-dimensional dynamical system, then it is Turing computable.

In sum, by defining personal identity as information transmission we can analyse personal identity by dynamical systems formalism. Such formalism can serve as the input to Turing machines. Glossing personal identity as non-branching, it can be modeled as a 1-dimensional dynamical system. These systems are decidable. So non-branching personal identity is decidable, which establishes *conclusion c*.

6 1-dimensional or *n*-dimensional personal identity

However, non-branching accounts encounter problems concerning transitivity which contradict the intuition of personal identity.

Take a person x at time t with a memory A. The same x at $t+1$ might still remember A, so $x_t = x_{t+1}$. Additionally, x_{t+1} also has a memory B. However, x at $t+2$ might not remember A, so $x_{t+2} \neq x_t$, but x_{t+2} does remember B, so $x_{t+2} = x_{t+1}$. This means that transitivity is broken, but x plausibly remains the same person.

Whilst various attempts have been made to save non-branching accounts,[29] the issue seems to either persist[30] or prompt a revision of tense which threatens to forfeit straightforward non-branching.[31]

To overcome these difficulties, Parfit has provided a distinction between two notions: identity and survival.[32] These lead to two relations:

- *Psychological connectedness*: an intransitive relation which requires the holding of direct psychological relations.
- *Psychological continuity*: a transitive relation which requires overlapping chains of direct psychological relations.

Parfit then terms that only connectedness is necessary for survival, whereas continuity is necessary for identity. He accepts that the two can come apart and notes that only survival, i.e. connectedness, is important. This entails that transitivity is not necessary for a person to survive. For example, if x_{t+2} does not remember A, then $x_{t+2} \neq x_t$, but x still survives as x_{t+2}. This resolves transitivity issues and preserves psychological continuity.

[27]Harrison 1969
[28]Lipton 1986
[29]See multiple occupancy, Lewis 1976, 21, and four dimensionalism, Sider 2001
[30]Parfit 1976, 94
[31]Lewis 1976, 26
[32]Parfit 1986, 204

However, Parfit's interpretation applies branching which has a consequence for dynamical systems. If continuity is broken a person branches onto a sub-branch of themselves, i.e. a different possibility of themselves. To illustrate Parfit describes a fission case as an additional problem for non-branching views.[33] A person x is split in two: x_1 and x_2. x is not identical to either x_1 or x_2, and $x_1 \neq x_2$ since they constitute independent persons. However, due to branching x is said to survive as both: x's timeline has branched into two sub-branches which ensure x stays connected to both x_1 and x_2. So, when two time-slices of a person exist on the same branch they are both connected and continuous, but when time-slices branch off they merely remain connected. For non-branching accounts this constitutes a fundamental problem: the person must cease to exist since only one continuous branch constitutes their being. However, branching accounts solve the tension by connectedness.

This entails that the dynamical system which models branching personal identity, or survival, has to be sufficiently robust to compute a person's total space of branches. However, since we take branches as the degrees of freedom for our system, 1-dimensional systems are not sufficient for this task. They can only model a single branch at all times. If two or more branches are the case, then the system fails. So the dynamical system must scale to n-dimensions to accommodate n number of branches.

7 Computing branching personal identity

n-dimensional dynamical systems exist in various forms which are either decidable or undecidable. However, I shall argue that any system which models branching personal identity must be chaotic. Any chaotic n-dimensional dynamical system is undecidable. So, branching personal identity is undecidable.

I shall argue as follows:

(d) Branching personal identity should be modeled as a chaotic n-dimensional dynamical system.

(e) Chaotic n-dimensional dynamical systems are undecidable.

(f) $\therefore$ Branching personal identity is undecidable.

Let us take each premise in turn.

7.1 Branching and chaotic dynamical systems

7.1.1 Branching and n-dimensional dynamical systems

Branching accounts term personal identity, or rather survival, as an event with n number of outcomes: between two information states of x, x can persist as $x_1, x_2, ..., x_n$ dependent upon branching. Personal identity must therefore be

[33] Parfit 1986, 253

modeled by n-dimensional dynamical systems, because the system has to be able to keep track of all possible branches.

Take Parfit's account as a paradigm example. Assume that at some time t, x splits in two: x_1 and x_2. At some time $t+1$, x_1 and x_2 stand to choose between two perfumes: {Chanel No. 5, Miss Dior}. Since x_1 and x_2 are independent persons they can commit to independent actions, so x_1 chooses Chanel No. 5, and x_2 Miss Dior. At this point the branches of x_1 and x_2 start to diverge. Say, for example, that y loves Chanel No. 5, but despises Miss Dior. So, x_1 and y live happily together ever after, but x_2 remains forever alone. This example is a proxy for a near infinite and continuous list of choices, so it is reasonable to conclude that x_1 and x_2's respective branches will diverge as their evolution progresses.

So, the two differing branches must be kept track of by the system modeling x's survival. Since the probability of every action is weighted by previous memories, and x_1 and x_2's memories diverge, the only option becomes an n-dimensional dynamical system.

In this way n-dimensional dynamical systems can accommodate Parfit's account. By this characterisation branching accounts can serve as inputs to a Turing machine.

7.1.2 Chaos

However, decidability is dependent on what type of n-dimensional dynamical system is deployed. I shall argue that only a chaotic system can record personal identity once it is elevated higher than 1-dimension.

Chaos is a mathematical property of dynamical systems. It entails that the system is highly sensitive to initial conditions. Had the initial conditions of the system been slightly different, the evolution of the system would have been drastically different.

Following Daveney,[34] I shall define a system S to be chaotic if it fulfills the following conditions:

1. *Denseness*: for each point x_1 in S there is a point x_2 that is arbitrarily close.
2. *Transitivity*: a dense system is automatically transitive because a dense orbit[35] must come arbitrarily close to all points in the system.
3. *Sensitivity*: if S has sensitive dependence on initial conditions, then we can always find a point x_2 within a distance of x_1 whose orbit eventually differs from that of x by a specified minimal distance ϵ.

7.1.3 Persons as chaotic dynamical systems

Let us apply chaos to personal identity and dynamical systems by analysing each chaos condition against branching and non-branching accounts.

[34] Devaney 1988

[35] An orbit is a sequence of points.

For (1), take the sub-branch for x_1. Since this sub-branch is a subset of all possible branches, there will always be an arbitrarily close correlation between the relevant set of sub-branches and all possible branches. This satisfies the *density* condition for branching accounts. Since non-branching accounts have no sub-branches, the condition fails to apply.

(2) is an immediate corollary of (1). *Transitivity* is further ensured in branching accounts under continuity, which non-branching accounts cannot satisfy.

Finally (3), since it is plausible to argue that any alteration in a person's total set of memories (their state ($|\psi\rangle$) can cause radical differences for the evolution of their branch, we can assume the *sensitivity* condition to hold for branching accounts. Since non-branching accounts have no sub-branches, there are no two branches to establish a measurement ϵ between. So the condition fails to apply.

This entails that branching personal identity fulfills conditions (1-3), whereas non-branching does not. Dynamical systems which model branching accounts must therefore be chaotic, whilst non-branching ones remain monotone.

We can illustrate this with Parfit's example again. At some time $t+2$, x_1 and x_2 stand to choose between some perfumes: {Chanel No. 5, Miss Dior, YSL Opium, Prada La Femme}. Since x_1's partner loves Chanel No. 5, x_1's probability distribution (ρ_{ψ_i}) is:

1. Chanel No. 5: 0.91
2. Miss Dior: 0.03
3. YSL Opium: 0.03
4. Prada La Femme: 0.03

Since x_2 is single, and still jealous of x_1 and y, x_2's probability distribution (ρ_{ψ_i}) is instead:

1. Chanel No. 5: 0.1
2. Miss Dior: 0.1
3. YSL Opium: 0.6
4. Prada La Femme: 0.3

The initial choice at $t+1$ is a perturbation ϵ_0 which continues to diverge x_1 and x_2's branches. Every action is weighted by previous memories, so the timelines of x_1 and x_2 will diverge exponentially as independent actions accumulate. It is therefore evident that ϵ_0 has a drastic effect on x's various branches. The evolution clearly deviates. Considering that ϵ_0 is a proxy which represents a near infinite list of perturbations: $\epsilon_1, \epsilon_2, ..., \epsilon_n$, it is clear that this dynamic system is chaotic.

By this I determine *premise d* fulfilled. Branching personal identity must be modeled as a chaotic n-dimensional dynamical system.

7.2 Chaotic n-dimensional dynamical systems are undecidable

Moving to *premise e* we see that personal identity, as established by *premise d*, is undecidable. Chaotic n-dimensional dynamical systems are undecidable because the relationships between the constants in their states are undecidable. This problem was initially posed by Penrose[36] regarding the Mandelbrot set, and was subsequently solved by Blum and Smale.[37] Arbieto and Matheus[38] imported these results and proved them for chaotic dynamical systems in general. So, if personal identity is a chaotic n-dimensional dynamical system, then it is not Turing computable.

In sum, n-dimensional dynamical systems can serve as inputs to Turing machines. However, glossing personal identity as branching, it is modeled by a chaotic dynamical system. These systems are undecidable. So branching personal identity is undecidable, which establishes *conclusion f.*

Since Parfit's account has been shown superior in overcoming both transitivity and fission issues, I determine his branching account to more accurately portray personal identity. From this it follows that personal identity is undecidable.

8 Conclusion

Our best conception of personal identity is branching, and if so, then it is not computable by deductive Turing machines. I have demonstrated this claim by: (i) characterising personal identity information theoretically which allows Turing computation by dynamical systems; (ii) determining that branching personal identity is superior to non-branching; and finally (iii) showing how branching accounts are undecidable. The outcome of this conclusion entails that deductive Turing machines, or any present artificial network, computing the person will at best achieve an approximation.

References

Arbieto A. Matheus, C. (2004). "Decidability of chaos for some families of dynamical systems". In: *Found. Comput. Math, 4, 269–275.*

Benthem, J. (2003). *Logical Dynamics of Information and Interaction.* Cambridge University Press.

Blum L. Smale, S. (1993). "The Godel incompleteness theorem and decidability over a ring". In: *From Topology to Computation: Proc. of the Smalefest, 321-339.*

Brown, B. T. et al. (2020). "Language Models are Few-Shot Learners". In: *NeurIPS 33, 1877-1901.*

[36] Penrose 1989, 177
[37] Blum 1993
[38] Arbieto 2004

Churchland, P. M. (1989). *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. MIT Press.

Demarest, H. (2015). "Fission May Kill You (But Not for the Reasons You Thought)". In: *Philosophy and Phenomenological Research, 93, 3, 565–582.*

Devaney, R. (1988). *A First Course in Chaotic Dynamical Systems.* A First Course in Chaotic Dynamical Systems.

Gelder, T. (1992). "What Might Cognition Be, If Not Computation?" In: *The Journal of Philosophy, 92, 7, 345-381.*

Harrison, M. A. (1969). *Lectures on Linear Sequential Machines.* Academic Press.

Kriegeskorte N. Douglas, K. P. (2018). "Cognitive Computational Neuroscience". In: *Nat Neurosci, 21, 1148–1160.*

Lewis, D. (1976). *The Identities of Persons.* University of California Press. Chap. Survival and Identity.

Lipton R. J. Kannan, R. (1986). "Polynomial-Time Algorithm for the Orbit Problem". In: *Journal ofthe Association for Computing Machinery, 33, 4, 808–821.*

Locke, J. (1694 (1975)). *An Essay Concerning Human Understanding.* Oxford University Press.

Olson, E. T. (1997). *The Human Animal: Personal Identity Without Psychology.* Oxford University Press.

Parfit, D. (1976). *The Identities of Persons.* University of California Press. Chap. Lewis, Perry, and what matters.

— (1986). *Reasons and Persons.* Oxford University Press.

Penrose, R. (1989). *The Emperor's New Mind.* Oxford University Press.

Ramsey, W. et al. (1990). "Connectionism, Eliminativism and The Future of Folk Psychology". In: *Philosophical Perspectives , 4, 499-533.*

Reid, T. (1785 (1969)). *Essays on the Intellectual Powers of Man.* MIT Press.

Schrittwieser, J. et al. (2020). "Mastering Atari, Go, chess and shogi by planning with a learned model". In: *Nature, 588, 604–609.*

Shannon, C. E. (1948). "A Mathematical Theory of Communication". In: *The Bell System Technical Journal, 27, 379–423.*

Sider, T. (2001). *Four Dimensionalism.* Oxford University Press.

Smith, B. C. (2019). *The Promise of Artificial Intelligence: Reckoning and Judgement.* The MIT Press.

Smith, P. (2013). *An Introduction to Gödel's Theorems.* Cambridge University Press.

Tanevski, J. et al. (2016). "Process-based design of dynamical biological systems". In: *Nature, Scientific Reports, 6, 34107.*

Thrun, S. et al. (2017). "Dermatologist-level classification of skin cancer with deep neural networks". In: *Nature 542, 115–118.*