Thought Curvature: An underivative hypothesis on the 'Supersymmetric Artificial Neural Network'

Jordan Micah Bennett

Abstract

Some supersymmetric Markov receptive sampling denoted $C^{\infty\pi}(R^{k\pi})$ for $k = m|n$ reasonably permits uniform symbols on some input space of form η, in the scope of empirical evidence pertaining to supersymmetry in the biological brain.[13] (See pseudocode)

Keywords: Supermathematics, Supermanifold, Supersymmetry, Lie Superalgebra

1 Overview

i. The aim is to contribute to the field of artificial general intelligence, often underlined as mankind's likely last invention.

ii. Machine learning often concerns constraining algorithms with respect to biological examples.

iii. Babies are great examples of **some non-trivial basis** for artificial general intelligence; babies are significant examples of biological baseis that are reasonably usable to inspire smart algorithms, especially in the aims of (i), regarding (ii). **Babies' brains** are fantastic measures of "**tabula rasa**"-like states, from which complicated abstractions are learnt into adulthood[29]; similar to how the recent breakthrough artificial intelligence program, "AlphaGo Zero", started out essentially "**blank**" beginning from random plays, up until it quickly **learnt** to become the planet's strongest go player today. (This quick outline highlights the **critical relevance** of ***games*** as necessary testbeds/algorithm **training** scenarios, in the aim of developing artificial general intelligence.)

"**Thought curvature**" subsumes the "**supermanifold hypothesis in deep learning**", while espousing the **importance** of considering biological constraints[23] in the aim of developing general machine learning models, pertinently, where babies' brains are observed to be pre-equipped with particular "**physics priors**", constituting specifically, the ability for babies to **intuitively know laws of physics**, while learning by reinforcement.[3]

It is palpable that the phrasing **“intuitively know laws of physics”** above, should **not be confused** for Nobel laureate or physics undergrad aligned babies that for example, write or understand physics papers/exams; instead, the aforesaid phrasing simply conveys that **babies’ brains are pre-baked with ways to naturally exercise physics based expectations w.r.t. interactions with objects in their world**, as indicated by Aimee Stahl and Lisa Feigenson.[3]

Outstandingly, the importance of **recognizing underlying causal physics laws in learning models** (although **not via supermanifolds**, as encoded in **Thought Curvature**), has recently been both demonstrated[21] and separately echoed by Deepmind[23], and of late, distinctly emphasized by Yoshua Bengio.[25]

2 Introduction

Deepmind’s atari q architecture[4] encompasses non-pooling convolutions, therein generating object shift sensitivity, whence the model maximizes some reward over said shifts together with separate changing states for each sampled **t** state; translation ***non-invariance***.

Separately, uetorch[24], encodes an object trajectory behaviour physics learner, particularly on pooling layers; translation ***invariance***.

It is non-abstrusely observable, that the childhood neocortical framework pre-encodes certain causal physical laws in the neurons[3], amalgamating in perceptual learning abstractions into non-childhood.

As such, it is perhaps exigent that non-invariant fabric composes in the invariant, pertinently in the margin of some asymptote entailing $\phi(x,\theta,\bar{\theta})^\top w$[1], therein engendering time-space complex optimal causal artificial construction.

3 Related work

There priorly existed translation variant/invariant manifold interaction paradigms[6][7], that effectively learn to disentangle varying factors. **However**, such models plausibly relent factors amidst **optimal**, causal - laws of physics arranged embeddings. (See **"A probable experiment"**)

4 A probable experiment: A Transverse Field Ising Spin (Super)–Hamiltonian Quantum Computation

Considering the Bessel aligned second-order linear damping equation: $\tilde{\phi} = (z + \frac{1}{\lambda})^{\frac{1}{2}}[C_1 I_{\frac{5}{2}}(\alpha(z + \frac{1}{\lambda})) + C_2 I_{-\frac{5}{2}}(\alpha(z + \frac{1}{\lambda}))]e^{\mu z}$[13] incorporating the travelling coordinate: $z = ax + by - ct$[13] whilst emerging by the isospectral factorization outcome: $f_\lambda = \frac{\lambda}{\lambda+1}$[13], **constrained** in dimensions $d = 2 + 1$[13], given that the $SO(n)$ group may eventuate in $SU(m|n)$ terms[17][36][45]; within the aforesaid **constraint**, the Hamiltonian operator: $-\sum_a \Gamma_a \sigma_a^x - \sum_a b_a \sigma_a^z - \sum_{a,b} w_{ab} \sigma_a^z \sigma_b^z$[14] is reasonably applicable in the quantum temporal difference horizon: $\pi(s1) \leftarrow argmax_a Q(s1, a)$[15] as a Super-Hamiltonian[16] in contrast.

Consequently, some odd operation of form $\{H \pm F, H \pm F\}1 = \pm 2QH, \{H + F, H - F\}1 = \{H \pm F, QH\}1 = \{QH, QH\}1 = 0$[16] subsuming $-\sum_a \Gamma_a \sigma_a^x - \sum_a b_a \sigma_a^z - \sum_{a,b} w_{ab} \sigma_a^z \sigma_b^z$[14] is theoretically absorbable in [15].

5 Thought Curvature - Annotation

i. **Supermanifold hypothesis in deep learning** describes some superspace bound structure/weights.

ii. Thought curvature designates the existence of some plausible π - Bellmanian[15], based on i; a sequence taking as parameters, the aforesaid supermanifold structure.

6 Supermanifold Hypothesis in Deep Learning

If any homeomorphic transition in some neighbourhood in an euclidean space R^n yields $\phi(x,\theta)^\top w$ for $w_i, \theta \in R^n$, then reasonably, some homeomorphic transition sequence in some euclidean superspace $C^\infty(R^{m|n})$ yields $\phi(x,\theta,\overline{\theta})^\top w$ for $x_i \in R^n$ while $w_i, \theta, \overline{\theta} \in R^{m|n}$.[3]

Pertinently, $R^{m|n} \to$ form $R^{2|1}$ for $d = 2 + 1$.[3]

7 Supermanifold Hypothesis in Deep Learning - Annotation

i. Deep learning entails $\phi(x,\theta)^\top w$[5], that denotes the input space x, and learnt representations θ.
ii. Deep learning underlines that coordinates or latent spaces in the manifold framework, are learnt features/representations, or **directions** that are **sparse configurations of coordinates**.
iii. Supermathematics entails $(x,\theta,\overline{\theta})$[9], that denotes some x valued coordinate distribution, and by extension, **directions** that **compact coordinates** via $\theta, \overline{\theta}$.

iv. As such, the aforesaid $(x,\theta,\overline{\theta})$, is subject to **coordinate transformation.**
v. Thereafter i, ii, iii, iv and [13], within the generalizable nature of euclidean space, reasonably effectuate $\phi(x,\theta,\overline{\theta})^\top w$.

8 Supermanifold Hypothesis in Deep learning - End notes

Some space (i.e. superspace) may persist, such that degrees of freedom of said space is inclined by some aggregation:

i. Some tensor sequence of priors. (i.e. eta… or any $\hat{p}_{\text{data}}$[5] constituting the **laws of physics**.)

ii. Some tensor sequence on the direction of (i), i.e. superfields of interactions in (i) terms.[28]

9 Thought Curvature - Limitations

Although **thought curvature** is minor particularly in its simple description (acquiescing SQCD[26]) in relation to Artificial General Intelligence, it **crucially** delineates that the math of supermanifolds is reasonably applicable in Deep Learning, imparting **that cutting edge Deep Learning work tends to consider boundaries in the biological brain**[23], while underscoring that biological brains can be **optimally** evaluated using supersymmetric operations.[13]

In broader words, thought curvature occurs on the following evidence:

1. **Manifolds** are in the regime of very **general algorithms**, that enable models **to learn many degrees of freedom** in latent space, (i.e. position, scale etc… where said degrees are observable as features of **physics** interactions) where transformations on points may represent for e.g., features of a particular object in pixel space, and transformations on said points or weights of an object are **disentangleable** or separable from those pertaining to other objects in latent space.[8][21][22]…
2. Given (1), and the generalizability of **euclidean space**, together with the instance that there persists **supersymmetric** measurements in biological brains, thought curvature predicates that **Supermathematics** or **Lie Superalgebras** (in **Supermanifolds**) may reasonably, empirically apply in **Deep Learning**, or some other named study of hierarchical learning in research.

10 A brief discussion on the significance of a Transverse Field Ising Spin (Super)-Hamiltonian reinforcement learning algorithm

The usage of **supersymmetric** operations is **imperatively** efficient, as such operations enable **deeply abstract representations** (as is naturally afforded by **symmetry group Lie Superalgebras**[26][27]), pertinently, in a **general, biologically tenable time-space complex optimal regime**.[13]

As such, said **deeply abstract representations** may **reasonably capture** certain **"physics priors"** (See page 1), with respect to the **laws of physics**.

11 An informal proof of the representation power gained by deeper abstractions of the "Supersymmetric Artificial Neural Network"

Machine learning non-trivially concerns the application of families of functions that **guarantee more and more variations** in weight space.

This means that machine learning researchers study what functions are best to transform the weights of the artificial neural network, such that the **weights learn** to represent **good values** for which **correct hypotheses or guesses** can be produced by the artificial neural network.

The "Supersymmetric Artificial Neural Network" (a core component in 'thought curvature') is yet another way to **represent richer values** in the weights of the model; because **supersymmetric values** can allow for **more information to be captured about the input space.** For example, supersymmetric systems can capture **potential-partner** signals, which is **beyond** the feature space of **magnitude** and **phase signals** learnt in **typical real valued neural nets** and **deep complex neural networks** respectively. As such, a brief historical progression of geometric solution spaces for varying neural network architectures follows:

1. An **optimal weight space** produced by **shallow** or low dimension **integer valued nodes or real valued artificial neural nets**, may have good weights that **lie for example**, **in one simple** $(\mathbb{Z}^n \ or \ \mathbb{R}^n - ordered)$ **cone per class/target group**. (This may guarantee **some variation**, but **not enough** for more sophisticated tasks of higher dimension)[30][35]

2. An **optimal weight space** produced by **deep** and high-dimension-absorbing **real valued artificial neural nets**, may have good weights that **lie in disentangleable** $(\mathbb{R}^n * \mathbb{R}^n - ordered)$ **manifolds per class/target group** convolved by the operator $*$, **instead** of the simpler regions per class/target group seen in item (1). (This may guarantee **more variation** in the weight space **than (1)**, leading to **better hypotheses or guesses**)[31]

3. An **optimal weight space** produced by **shallow** but high dimension-absorbing **complex valued artificial neural nets**, may have good weights that **lie in multiple** $(\mathbb{C}^n - ordered)$ **sectors per class/target group**, **instead** of the real regions per class/target group seen amongst the prior items. (This may guarantee **more variation** of the weight space **than the previous items,** by **learning additional features**, in the "phase space". This also leads to **better hypotheses/guesses**)[32]

4. An **optimal weight space** produced by **deep** or high dimension-absorbing **complex valued artificial neural nets**, may have good weights that **lie in chi distribution bound,** $(\mathbb{C}^n * \mathbb{C}^n - ordered)$ **rayleigh space per class/target group** convolved by the operator $*$, **instead** of the simpler sectors/regions per class/target group seen amongst the previous items. (This may guarantee **more variation** of the weight space **than the prior items**, by learning **phase space representations**, and **by extension**,

strengthen these representations via convolutional residual blocks. This also leads to **better hypotheses/guesses)**[33]

5. The "Supersymmetric Artificial Neural Network" operable on high dimensional data, may reasonably generate **good weights** that **lie in disentangleable** $(C^{\infty}(R^{m|n}) - ordered)$ **supermanifolds per class/target group**, **instead** of the solution geometries seen **in the prior items above**. Supersymmetric values can encode **rich partner-potential delimited features beyond the phase space of (4)** in accordance with **cognitive biological space**[13], **where (4) lacks the partner potential formulation** describable in Supersymmetric embedding.[34]

12 Pseudocode for the "Supersymmetric Artificial Neural Network"

Following, is another view of "solution geometry" history, which may promote a clear way to view the reasoning behind the subsequent pseudocode sequence:

1. There has been a clear progression of "**solution geometries**", ranging from those of the ancient **Perceptron**[30] to **complex valued neural nets** [33], **grassmann manifold artificial neural networks**[39] or **unitaryRNNs**.[38][40][44] These models may be denoted by $\phi(x, \theta)^{\top} w$ parameterized by θ, expressible as geometrical groups ranging from orthogonal[17] to special unitary group[46] based: $SO(n)$ to $SU(n)$…, and they got **better at representing input data i.e. representing richer weights**, thus the learning models generated better hypotheses or guesses.
2. By "**solution geometry**" I mean simply the class of regions where an algorithm's weights may lie, when generating those weights to do some task.
3. As such, if one follows **cognitive science**, one would know that biological brains may be measured in terms of **supersymmetric** operations. (Perez et al, "Supersymmetry at brain scale")
4. These supersymmetric biological brain representations can be represented by supercharge[37] compatible special unitary notation $SU(m|n)$, or $\phi(x, \theta, \bar{\theta})^{\top} w$ parameterized by $\theta, \bar{\theta}$ [34], which are supersymmetric directions, unlike θ seen in item (1). Notably, Supersymmetric values can **encode or represent more information** than the prior classes seen in (1), in terms of "partner potential" signals for example.
5. So, state of the art machine learning work forming $U(n)$ or $SU(n)$ based solution geometries, although ***non-supersymmetric***, are **already** in the **family** of ***supersymmetric*** **solution geometries** that may be observed as occurring in biological brain or $SU(m|n)$ supergroup **representation**.

I call an "***Edward Witten/String theory powered artificial neural network***", 'simply' an artificial neural network that learns supersymmetric weights.

Looking at the above progression of 'solution geometries'; going from $SO(n)$ [30] representation to $SU(n)$[38][44] representation has guaranteed richer and richer representations in weight space of the artificial neural network, and hence better and better hypotheses were generatable.

It is perhaps only then reasonable to look to $SU(m|n)$ representation, i.e. the "***Edward Witten/String theory powered artificial neural network***" ("***Supersymmetric Artificial Neural Network***").

To construct an "***Edward Witten/String theory powered artificial neural network***", it may be feasible to start with a **grassmann manifold artificial neural network** then **generate 'charts'**[43] until **scenarios occur**[34] where the "***Edward Witten/String theory powered artificial neural network***" is achieved in the following way:

Pseudocode:

a. Initialize input Supercharge[37] compatible special unitary matrix $SU(m|n)$. [45] (This is the atlas seen in b.)

b. Compute ∇C w.r.t. to $SU(m|n)$, where C is some cost manifold.

- Weight space is reasonably some Kähler potential like form: $K(\phi, \phi^*)$, obtained on some initial projective space CP^{n-1}.[42]
- It is feasible that CP^{n-1} (a C^∞ bound **atlas**) may be obtained from charts of **grassmann manifold networks**[39] where there exists some invertible submatrix entailing matrix $A \in \phi_i(U_i \cap U_j)$ for $U_i = \pi(V_i)$ where π is a submersion mapping enabling some differentiable grassmann manifold $GF_{k,n}$, and $V_i = u \in R^{n\times k}: \det(u_i) \neq 0$.[43]

c. Parameterize $SU(m|n)$ in $-\nabla C$ terms, by **Darboux transformation**.[41]

d. Repeat until convergence.

13 Thought Curvature - Experimentation considerations for Supersymmetric Reinforcement learning

Pertinently, an ***initial degree*** of the **(Super-) Hamiltonian**[16] structure required by **thought curvature** shall require a quite scalable scheme, such as some **boson sampling**[20] aligned range, in conjunction with **supersymmetric space**.[11] This scheme is approachable on the scale of **42 qubits**[19], or a 42 qubit = 32 x 2^{12} gb = 131,072 gb ram configuration for simple task/circuit tests.

More testing is required to determine the **model's** feasibility, and unravel $\hat{p}_{\text{data}}$ (training sample) types applicable to the **model**.

References

1. Jordan Bennett "Supermanifold Hypothesis in Deep Learning", 2016.
2. Brenden M. Lake, Tomer D. Ullman et al. "Building machines that learn and think like people", 2016.
3. Aimee E. Stahl, Lisa Feigenson, “Observing the unexpected enhances infants’ learning and exploration”, 2015.
4. Volodymyr Mnih et al, “Playing Atari with Deep Reinforcement Learning”, 2016.
5. Yoshua Bengio et al, “Deep Learning Book”, 2016.
6. Kihyuk Sohn, Yuting Zhang et al, “Learning to Disentangle Factors of Variation with Manifold Interaction”, 2014.
7. Yoshua Bengio et al, “Disentangling Factors of Variation via Generative Entangling”, 2012.
8. Christopher Olah, "Neural Networks, Manifolds, and Topology", 2014.
9. Wikipedia, “Supermanifolds”.
10. Wikipedia, “Differentiable Manifolds”.
11. Ncatlab, “Euclidean supermanifold”.
12. Christopher Lu, “Artificial neural network for behaviour learning from meso-scale simulations, application to multi-scale multi-material flows”, 2010.
13. P´erez et al. “Supersymmetric methods in the travelling variable: inside neurons and at the brain scale”, 2007.
14. Mohammad H. Amin, Evgeny Andriyash et al. “Quantum Boltzmann Machine”, 2016.

15. Daniel Crawford, Anna Levit et al. “Reinforcement Learning Using Quantum Boltzmann Machines”, 2016.
16. Armen Nersessian “Elements of (super-)Hamiltonian Formalism”, 2005.
17. Wikipedia “Orthogonal Groups”.
18. Wikipedia “Temporal Difference Learning’.
19. Alex Neville, Anthony Laing et al. “No imminent quantum supremacy by boson sampling”, 2017
20. Wikipedia “Boson sampling”
21. Irina Higgins, Loic Matthey et al. “Early Visual Concept Learning with Unsupervised Deep Learning”, 2016
22. Ben Poole, Subhaneil Lahiri et al. “Exponential expressivity in deep neural networks through transient chaos”, 2016
23. Demis Hassabis, Dharshan Kumaran et al. “Neuroscience-Inspired Artificial Intelligence”, 2017
24. Adam Lerer, Sam Gross et al. “Learning Physical Intuition of Block Towers by Example”, 2016
25. Yoshua Bengio, “The Consciousness Prior”, 2017
26. David Poland, David Simmons-Duffin, “N = 1 SQCD and the Transverse Field Ising Model”, 2011
27. Wikipedia “Symmetry Group”
28. Jelle Hartong, et al. “The supersymmetric tensor hierarchy of N = 1, d = 4 supergravity” 2009
29. Bruno B Averbeck, Vincent D Costa "Dopamine and reward prediction errors", 2016

30. Wikipedia “Perceptron”

31. Wikipedia “Deep Learning#Deep Neural Networks”

32. Tariq Rashid “Complex Valued Neural Networks - Experiments”, 2016

33. Chiheb Trabelsi, Olexa Bilaniuk et al. “Deep Complex Networks”, 2017

34. Wikipedia “Supersymmetry”

35. Wikipedia “Artificial Neural Network#Hebbian learning”

36. Diederik Aerts, Marek Czachor et al “Quantum Aspects of Semantic Analysis and Symbolic Artificial Intelligence”, 2004

37. Wikipedia “Supercharge”

38. Simone Fiori “A study on neural learning on manifold foliations: the case of the Lie group SU(3)”, 2008

39. Zhiwu Huang, Jiqing Wu et. al. “Building Deep Networks on Grassmann Manifolds”, 2017

40. Martin Arjovsky, Amar Shah et. al. “Unitary Evolution Recurrent Neural Networks”, 2015

41. Encyclopedia of math “Darboux transformation”

42. Kiyoshi Higashijima, Muneto Nitta “Supersymmetric Nonlinear Sigma Models”, 2000

43. Math wisc edu “The Grassmann Manifold”

44. Li Jing, Yichen Shen, “Tunable Efficient Unitary Neural Networks (EUNN) and their application to RNNs”, 2017

45. Wikipedia “Supergroup_(physics)”

46. Wikipedia “Special unitary group”