Beren Millidge

Email: beren@millidge.name
LinkedIn:
www.linkedin.com/in/beren-millidge-377065142/
GitHub: github.com/BerenMillidge

Education

University of Edinburgh — Edinburgh
Ph.D. in Machine Learning and Computational Neuroscience, supervised by Richard Shillcock — 2017–2021

University of Edinburgh — Edinburgh
MSc. in Artificial Intelligence, **Distinction** — 2016–2017

University of Oxford — Oxford
BSc. in Psychology, Philosophy, and Linguistics, **First Class Honours** — 2013–2016

Experience

Zyphra — Palo Alto, USA
Cofounder and Chief Scientist — July 2023 –present
I lead the technical teams at Zyphra.

Apollo Research — London, UK
Cofounder — May 2023 –July 2023
One of four cofounders. Aided in initial fundraising, organizational setup, direction-setting, and hiring of initial team. Left to start Zyphra.

Conjecture — London, UK
Head of Research — September 2022 –May 2023
Managed a research team of 4 full time researchers and 2 interns. Technical AI alignment research, primarily interpretability on large language models.

Verses Research Lab — Los Angeles, CA
Senior Research Scientist — February 2022 –September 2022
Consulted on applied AI problems involving vision, robotics, and logistics

University of Oxford — Oxford,UK
Postdoctoral Researcher — April 2021 –September 2022
Working with Rafal Bogacz on predictive coding, understanding learning and backprop in the brain, and unifying machine learning and neuroscience.

University of Sussex — Brighton,UK
Visiting Fellow — January 2020 –Current
Worked closely with Christopher Buckley and Anil Seth on a wide range of machine learning and computational neuroscience projects.

Areas of Expertise

- **Machine Learning** Model-Based and Model-Free Reinforcement Learning, Large Language models, Transformer and SSM architectures, Predictive Coding
- **Machine Learning Libraries** Pytorch, Jax, Flux.jl
- **Bayesian Inference** Variational Inference, Graphical Models, MCMC, (numpyro, Edward, Turing.jl)
- **Statistical Analysis** ANOVA, LMER
- **Web Development** Django, Flask, Nodejs, Express.js, React, Vue.js, Shiny (R)

Languages

- **Highly Experienced** Python, Javascript, Julia
- **Proficient** C++, C, Java, Rust, HTML, CSS, R
- **Conversant** CUDA, Ruby, Haskell, Elm, Bash, PHP, Typescript, MATLAB

Publications

[1] N. Alonso and **B. Millidge**, "Mixture-of-pageranks: Replacing long-context with real-time, sparse graphrag", *arXiv preprint arXiv:2412.06078*, 2024.

[2] N. Alonso, T. Figliolia, A. Ndirango, and **B. Millidge**, "Toward conversational agents with context and time sensitive long-term memory", *arXiv preprint arXiv:2406.00057*, 2024.

[3] Q. Anthony, Y. Tokpanov, P. Glorioso, and **B. Millidge**, "Blackmamba: Mixture of experts for state-space models", *arXiv preprint arXiv:2402.01771*, 2024.

[4] C. L. Buckley, T. Lewens, M. Levin, **B. Millidge**, A. Tschantz, and R. A. Watson, "Natural induction: Spontaneous adaptive organisation without natural selection", *Entropy*, vol. 26, no. 9, p. 765, 2024.

[5] K. J. Friston, M. J. Ramstead, A. B. Kiefer, A. Tschantz, C. L. Buckley, M. Albarracin, R. J. Pitliya, C. Heins, B. Klein, **B. Millidge**, *et al.*, "Designing ecosystems of intelligence from first principles", *Collective Intelligence*, vol. 3, no. 1, p. 26 339 137 231 222 481, 2024.

[6] P. Glorioso, Q. Anthony, Y. Tokpanov, A. Golubeva, V. Shyam, J. Whittington, J. Pilault, and **B. Millidge**, "The zamba2 suite: Technical report", *arXiv preprint arXiv:2411.15242*, 2024.

[7] P. Glorioso, Q. Anthony, Y. Tokpanov, J. Whittington, J. Pilault, A. Ibrahim, and **B. Millidge**, "Zamba: A compact 7b ssm hybrid model", *arXiv preprint arXiv:2405.16712*, 2024.

[8] C. Heins, **B. Millidge**, L. Da Costa, R. P. Mann, K. J. Friston, and I. D. Couzin, "Collective behavior from surprise minimization", *Proceedings of the National Academy of Sciences*, vol. 121, no. 17, e2320239121, 2024.

[9] **B. Millidge**, Y. Song, A. Lak, M. E. Walton, and R. Bogacz, "Reward bases: A simple mechanism for adaptive acquisition of multiple reward types", *PLOS Computational Biology*, vol. 20, no. 11, e1012580, 2024.

[10] **B. Millidge**, M. Tang, M. Osanlouy, N. S. Harper, and R. Bogacz, "Predictive coding networks for temporal prediction", *PLOS Computational Biology*, vol. 20, no. 4, e1011183, 2024.

[11] A. Ororbia, A. Mali, A. Kohan, **B. Millidge**, and T. Salvatori, "A review of neuroscience-inspired machine learning", *arXiv preprint arXiv:2403.18929*, 2024.

[12] V. Shyam, J. Pilault, E. Shepperd, Q. Anthony, and **B. Millidge**, "Tree attention: Topology-aware decoding for long-context attention on gpu clusters", *arXiv preprint arXiv:2408.04093*, 2024.

[13] Y. Song, **B. Millidge**, T. Salvatori, T. Lukasiewicz, Z. Xu, and R. Bogacz, "Inferring neural activity before plasticity as a foundation for learning beyond backpropagation", *Nature neuroscience*, vol. 27, no. 2, pp. 348–358, 2024.

[14] Y. Tokpanov, **B. Millidge**, P. Glorioso, J. Pilault, A. Ibrahim, J. Whittington, and Q. Anthony, "Zyda: A 1.3 t dataset for open language modeling", *arXiv preprint arXiv:2406.01981*, 2024.

[15] M. Aguilera, **B. Millidge**, A. Tschantz, and C. L. Buckley, "From the free energy principle to a confederation of bayesian mechanics: Reply to comments on" how particular is the physics of the free energy principle?"", *Physics of life reviews*, vol. 44, pp. 270–275, 2023.

[16] M. De Llanza Varona, C. Buckley, and **B. Millidge**, "Exploring action-centric representations through the lens of rate-distortion theory", in *International Workshop on Active Inference*, Springer, 2023, pp. 189–203.

[17] M. J. Ramstead, D. A. Sakthivadivel, C. Heins, M. Koudahl, **B. Millidge**, L. Da Costa, B. Klein, and K. J. Friston, "On bayesian mechanics: A physics of and by beliefs", *Interface Focus*, vol. 13, no. 3, p. 20 220 029, 2023.

[18] T. Salvatori, **B. Millidge**, Y. Song, R. Bogcaz, and T. Lukasiewicz, "Associative memories in the feature space", in *ECAI 2023*, IOS Press, 2023, pp. 2065–2072.

[19] T. Salvatori, L. Pinchetti, A. M'Charrak, **B. Millidge**, and T. Lukasiewicz, "Causal inference via predictive coding", *arXiv preprint arXiv:2306.15479*, 2023.

[20] M. Tang, T. Salvatori, **B. Millidge**, Y. Song, T. Lukasiewicz, and R. Bogacz, "Recurrent predictive coding models for associative memory employing covariance learning", *PLoS computational biology*, vol. 19, no. 4, e1010719, 2023.

[21] A. Tscshantz, **B. Millidge**, A. K. Seth, and C. L. Buckley, "Hybrid predictive coding: Inferring, fast and slow", *PLoS Computational Biology*, vol. 19, no. 8, e1011280, 2023.

[22] N. Alonso, **B. Millidge**, J. Krichmar, and E. Neftci, "A theoretical framework for inference learning", *arXiv preprint arXiv:2206.00164*, 2022.

[23] S. Black, L. Sharkey, L. Grinsztajn, E. Winsor, D. Braun, J. Merizian, K. Parker, C. R. Guevara, **B. Millidge**, G. Alfour, *et al.*, "Interpreting neural networks through the polytope lens", *arXiv preprint arXiv:2211.12312*, 2022.

[24] C. Heins, **B. Millidge**, D. Demekas, B. Klein, K. Friston, I. Couzin, and A. Tschantz, "Pymdp: A python library for active inference in discrete state spaces", *arXiv preprint arXiv:2201.03904*, 2022.

[25] A. B. Kiefer, **B. Millidge**, A. Tschantz, and C. L. Buckley, "Capsule networks as generative models", in *International Workshop on Active Inference*, Springer, 2022, pp. 192–209.

[26] P. F. Kinghorn, **B. Millidge**, and C. L. Buckley, "Preventing deterioration of classification accuracy in predictive coding networks", in *International Workshop on Active Inference*, Springer, 2022, pp. 1–15.

[27] **B. Millidge** and S. Black, "The singular value decompositions of transformer weight matrices are highly interpretable", in *AI Alignment Forum*, 2022, p. 17.

[28] **B. Millidge** and C. L. Buckley, "Active inference successor representations", in *International Workshop on Active Inference*, Springer, 2022, pp. 151–161.

[29] **B. Millidge**, T. Salvatori, Y. Song, R. Bogacz, and T. Lukasiewicz, "Predictive coding: Towards a future of deep learning beyond backpropagation?", *arXiv preprint arXiv:2202.09467*, 2022.

[30] **B. Millidge**, T. Salvatori, Y. Song, T. Lukasiewicz, and R. Bogacz, "Universal hopfield networks: A general framework for single-shot associative memory models", *arXiv preprint arXiv:2202.04557*, 2022.

[31] **B. Millidge**, Y. Song, T. Salvatori, T. Lukasiewicz, and R. Bogacz, "A theoretical framework for inference and learning in predictive coding networks", *arXiv preprint arXiv:2207.12316*, 2022.

[32] **B. Millidge**, Y. Song, T. Salvatori, T. Lukasiewicz, and R. Bogacz, "Backpropagation at the infinitesimal inference limit of energy-based models: Unifying predictive coding, equilibrium propagation, and contrastive hebbian learning", *arXiv preprint arXiv:2206.02629*, 2022.

[33] **B. Millidge**, M. Walton, and R. Bogacz, "Reward bases: Instantaneous reward revaluation with temporal difference learning", *bioRxiv*, 2022.

[34] T. Salvatori, L. Pinchetti, **B. Millidge**, Y. Song, R. Bogacz, and T. Lukasiewicz, "Learning on arbitrary graph topologies via predictive coding", *arXiv preprint arXiv:2201.13180*, 2022.

[35] T. Salvatori, Y. Song, **B. Millidge**, Z. Xu, L. Sha, C. Emde, R. Bogacz, and T. Lukasiewicz, "Incremental predictive coding: A parallel and fully automatic learning algorithm", *arXiv preprint arXiv:2212.00720*, 2022.

[36] L. Sharkey, D. Braun, and **B. Millidge**, "Taking features out of superposition with sparse autoencoders", in *AI Alignment Forum*, vol. 6, 2022, pp. 12–13.

[37] Y. Song, B. G. Millidge, T. Salvatori, T. Lukasiewicz, Z. Xu, and R. Bogacz, "Inferring neural activity before plasticity: A foundation for learning beyond backpropagation", *bioRxiv*, 2022.

[38] A. Tschantz, **B. Millidge**, A. K. Seth, and C. L. Buckley, "Hybrid predictive coding: Inferring, fast and slow", *arXiv preprint arXiv:2204.02169*, 2022.

[39] M. Aguilera, **B. Millidge**, A. Tschantz, and C. L. Buckley, "How particular is the physics of the free energy principle?", *arXiv preprint arXiv:2105.11203*, 2021.

[40] P. F. Kinghorn, **B. Millidge**, and C. L. Buckley, "Habitual and reflective control in hierarchical predictive coding", *arXiv preprint arXiv:2109.00866*, 2021.

[41] P. Lanillos, C. Meo, C. Pezzato, A. A. Meera, M. Baioumy, W. Ohata, A. Tschantz, **B. Millidge**, M. Wisse, C. L. Buckley, *et al.*, "Active inference in robotics and artificial agents: Survey and challenges", *arXiv preprint arXiv:2112.01871*, 2021.

[42] **B. Millidge**, "Applications of the free energy principle to machine learning and neuroscience", *arXiv preprint arXiv:2107.00140*, 2021.

[43] **B. Millidge**, "Towards a mathematical theory of abstraction", *arXiv preprint arXiv:2106.01826*, 2021.

[44] **B. Millidge**, A. Seth, and C. L. Buckley, "Predictive coding: A theoretical and experimental review", *arXiv preprint arXiv:2107.12979*, 2021.

[45] **B. Millidge**, A. Tschantz, A. Seth, and C. Buckley, "Neural kalman filtering", *arXiv preprint arXiv:2102.10021*, 2021.

[46] **B. Millidge**, A. Tschantz, A. Seth, and C. Buckley, "Understanding the origin of information-seeking exploration in probabilistic objectives for control", *arXiv preprint arXiv:2103.06859*, 2021.

[47] A. D. Noel, C. van Hoof, and **B. Millidge**, "Online reinforcement learning with sparse rewards through an active inference capsule", *arXiv preprint arXiv:2106.02390*, 2021.

[48] **B. Millidge**, A. Tschantz, and C. L. Buckley, "Predictive coding approximates backprop along arbitrary computation graphs", *arXiv preprint arXiv:2006.04182; Submitted to ICLR 2021*, 2020.

[49] **B. Millidge**, A. Tschantz, and C. L. Buckley, "Whence the expected free energy?", *Neural Computation*, 2020.

[50] **B. Millidge**, A. Tschantz, C. L. Buckley, and A. Seth, "Activation relaxation: A local dynamical approximation to backpropagation in the brain", *arXiv preprint arXiv:2009.05359; submitted to ICLR 2021*, 2020.

[51] **B. Millidge**, A. Tschantz, C. L. Buckley, and A. Seth, "Investigating the scalability and biological plausibility of the activation relaxation algorithm", *arXiv preprint arXiv:2009.05359; submitted to NeurIPS 2020 workshop, Beyond Backpropagation in the Brain*, 2020.

[52] **B. Millidge**, A. Tschantz, A. Seth, and C. L. Buckley, "Relaxing the constraints on predictive coding models", *arXiv preprint arXiv:2010.01047; submitted to Neural Networks*, 2020.

[53] **B. Millidge**, A. Tschantz, A. K. Seth, and C. L. Buckley, "On the relationship between active inference and control as inference", *IEEE IWAI Workshop on Active Inference*, 2020.

[54] **B. Millidge**, A. Tschantz, A. K. Seth, and C. L. Buckley, "Reinforcement learning as iterative and amortised inference", *arXiv preprint arXiv:2006.10524*, 2020.

[55] A. Seth, **B. Millidge**, C. L. Buckley, and A. Tschantz, "Curious inferences: Reply to sun & firestone on the dark room problem", *Trends in Cognitive Science*, 2020.

[56] A. Tschantz, **B. Millidge**, A. K. Seth, and C. L. Buckley, "Control as hybrid inference", *ICML 2020 Workshop on Theoretical Foundations of RL*, 2020.

[57] A. Tschantz*, **B Millidge***, A. K. Seth, and C. L. Buckley, "Reinforcement learning through active inference", *ICML 2020 Workshop Bridging AI and Cognitive Science*, 2020.

[58] **B. Millidge**, "Combining active inference and hierarchical predictive coding: A tutorial introduction and case study", *PsyArxiv; Submitted to Cognition*, 2019.

[59] **B. Millidge**, "Deep active inference as variational policy gradients", *Journal of Mathematical Psychology*, vol. 96, p. 102 348, 2019.

[60] **B. Millidge**, "Fixational eye movements: Data augmentation for the brain?", *PsyArxiv*, 2019.

[61] **B. Millidge**, "Implementing predictive processing and active inference: Preliminary steps and results", *PsyArxiv*, 2019.

[62] R. Shillcock, **B. Millidge**, and A. Ravignani, "Exploring infant vocal imitation in tadarida brasiliensis mexicana", in *Neurobiology of Speech and Language*, 2019, pp. 36–37.

[63] **B. Millidge** and R. Shillcock, "A predictive processing account of bottom-up visual saliency using cross-predicting autoencoders", *PsyArxiv*, 2018.

Awards

- Best Dissertation in Artificial Intelligence Award, University of Edinburgh 2017
- Highest Performance in Prelims Linguistics, University of Oxford 2014