Neuromorphic cognition and embodied AI: from Neuroscience to Robotics and Back

intel labs

Yulia Sandamirskaya
Neuromorphic computing Lab, Intel Labs (Munich)

Shanghai Lecture, 01. Dec 2022

Algorithms for Artificial Intelligence Today

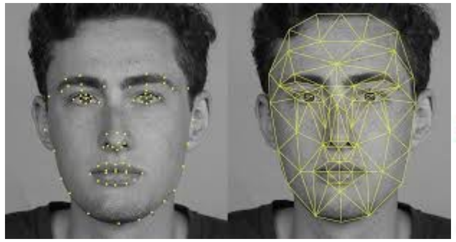


Image (vide

Label, ROI,
ommand

Sound sample

Text, picture

Algorithms for Artificial Intelligence Today and Tomorrow

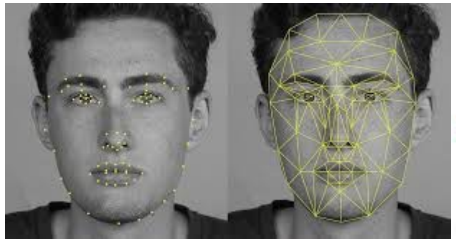
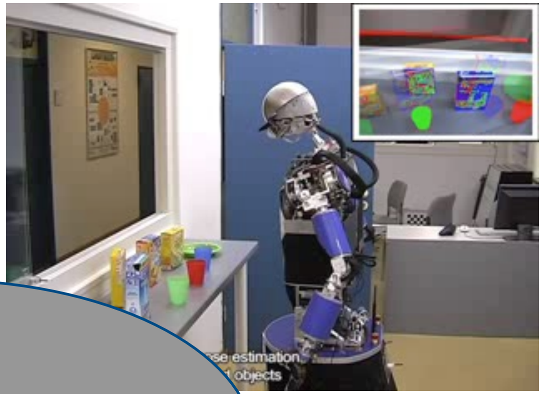




Image (vide

Label, ROI,
mmand

Sound
sequence

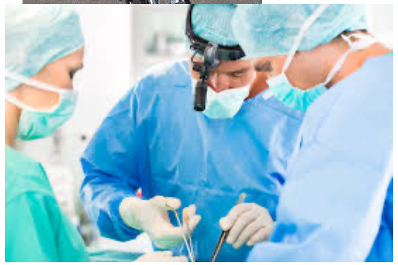
ext, picture

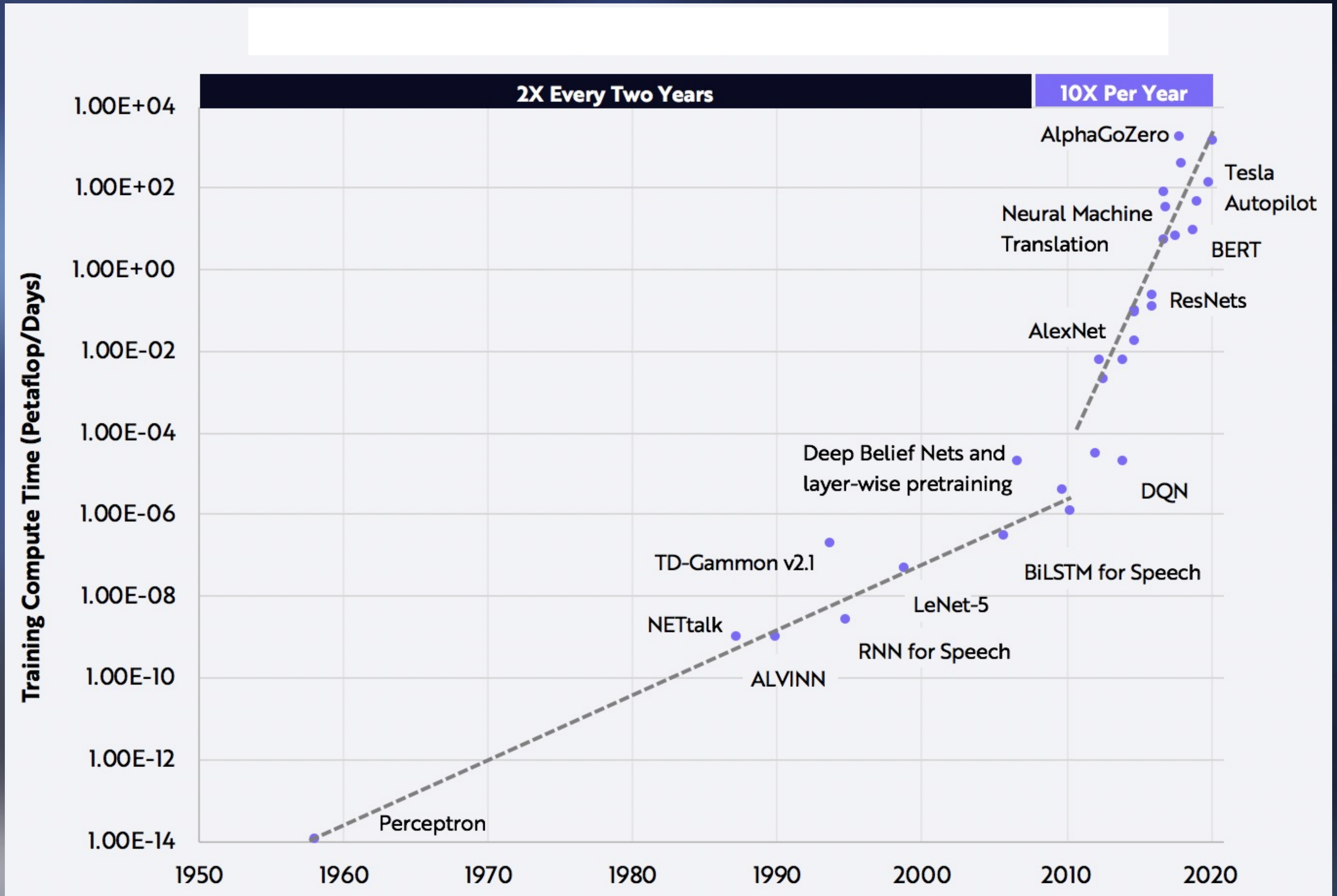
?







The Gains of Deep Learning Come at Increasing Cost



300,000x increase in required training computation over 6 years vs **8x** provided by the Moore's Law

? Efficiency
? Transparency, robustness
? Adaptivity

Source: ARK Investment Management LLC, "AI and Compute." OpenAI, https://arkinv.st/2ZOH2Rr.
https://laptrinhx.com/news/the-cost-of-ai-training-is-improving-at-50x-the-speed-of-moore-s-law-why-it-s-still-early-days-for-ai-jY1BQeq/

Biological intelligence

- 1g brain, 1M neurons, 1mW
- Navigates and learns in unknown environments "on the fly"

- 2.2g brain, 10 M neurons, 50 mW
- Navigates and learns "on the fly"
- Can learn words
- Can learn to manipulate objects

- 1000g brain, 100 B neurons, 20 W
- Can do amazing things

Biological brains:

Adaptive
Flexible
Fast
Precise
Efficient

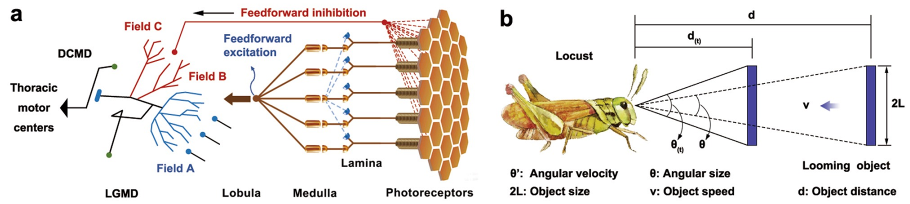
Can deal with real-world complexity

Learn new tasks

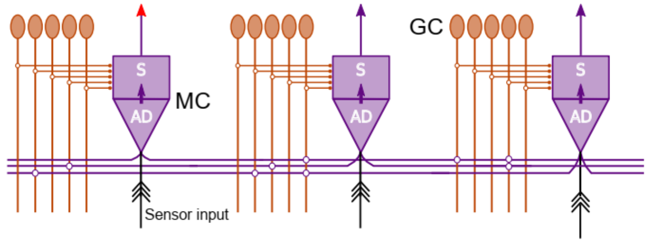
"Cognitive"

What can we learn from biology?

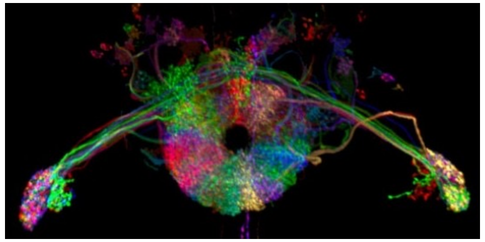
1. Diversity of neuron types, connectivity motives, network structures and topologies



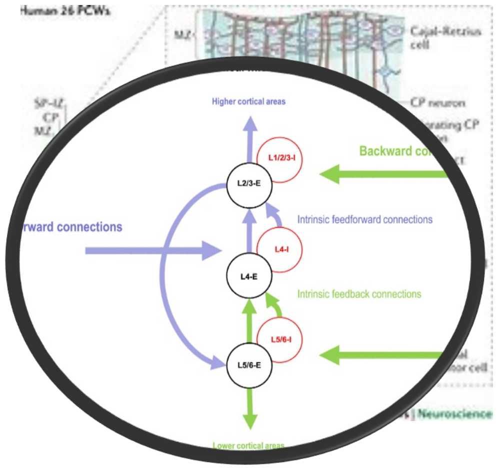
Locust's Giant Motion Detector neuron (LGMD)



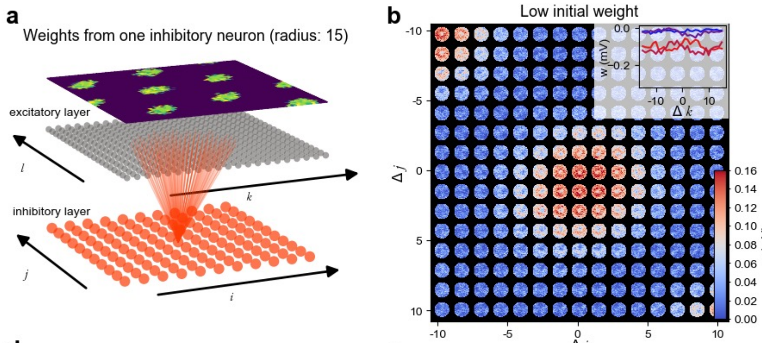
Olfactory circuits



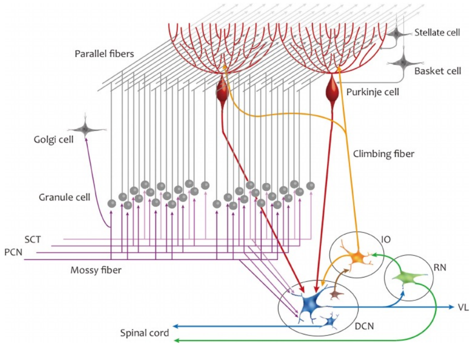
Fly's head direction circuit



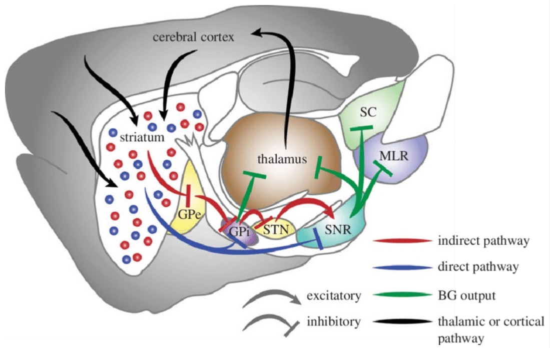
Neocortical layers



Grid cell, hippocampal circuits



Cerebellar architecture



Basal ganglia

2. A lot of predetermined structure augmented with continual learning and plasticity

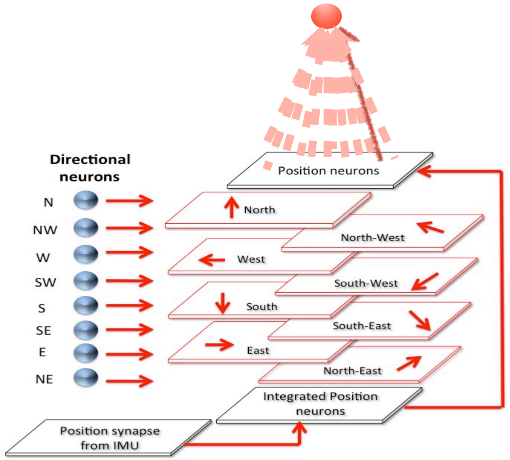
How can we learn from biology?

➢ We can learn specific neural circuits for different tasks

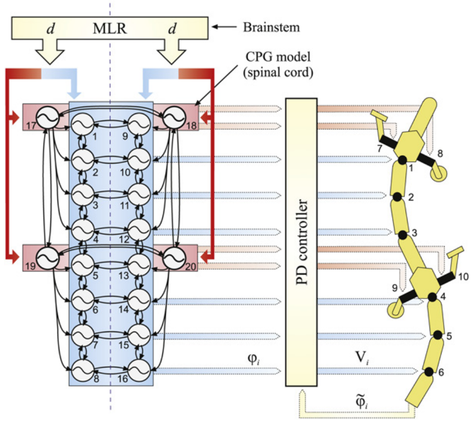
1) Sensing (LGMD, divergence-based landing)
2) Navigation (hippocampal circuits, RatSLAM)
3) CPGs for locomotion

1)

2)



3)



➢ We can learn architectural principles

1) Statful computing; states dynamically stabilitized
2) Loops (predictions, consitency checks)
3) (Autonomous) learning principles

Collection of "solutions" Algorithm, Mechanisms) "Bio" ⟷ Principles, laws, building blocks "Phys"

Technology "Eng"

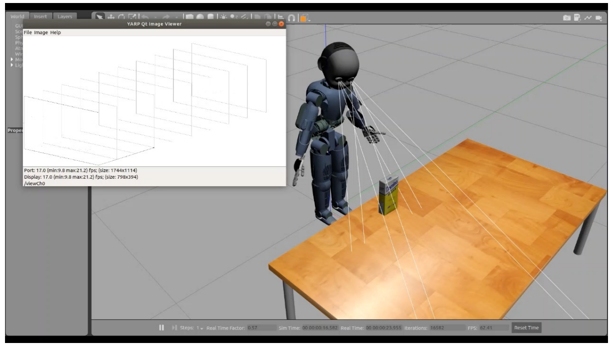
1) Salt, L., Indiveri, G., & Sandamirskaya, Y. (2017, May). Obstacle avoidance with LGMD neuron: towards a neuromorphic UAV implementation. In *2017 IEEE International Symposium on Circuits and Systems (ISCAS)* (pp. 1-4). IEEE.

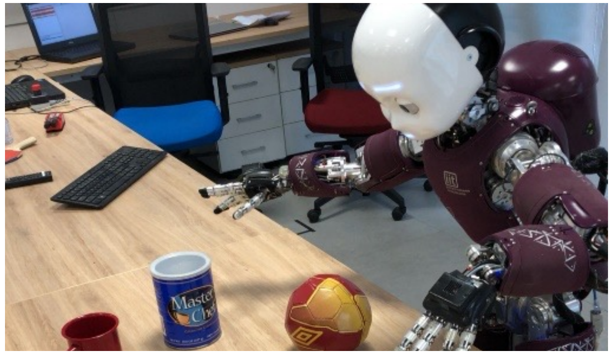
2) Kreiser, R., Renner, A., Sandamirskaya, Y., & Pienroj, P. (2018, October). Pose estimation and map formation with spiking neural networks: towards neuromorphic SLAM. In *IROS* (pp. 2159-2166). IEEE.

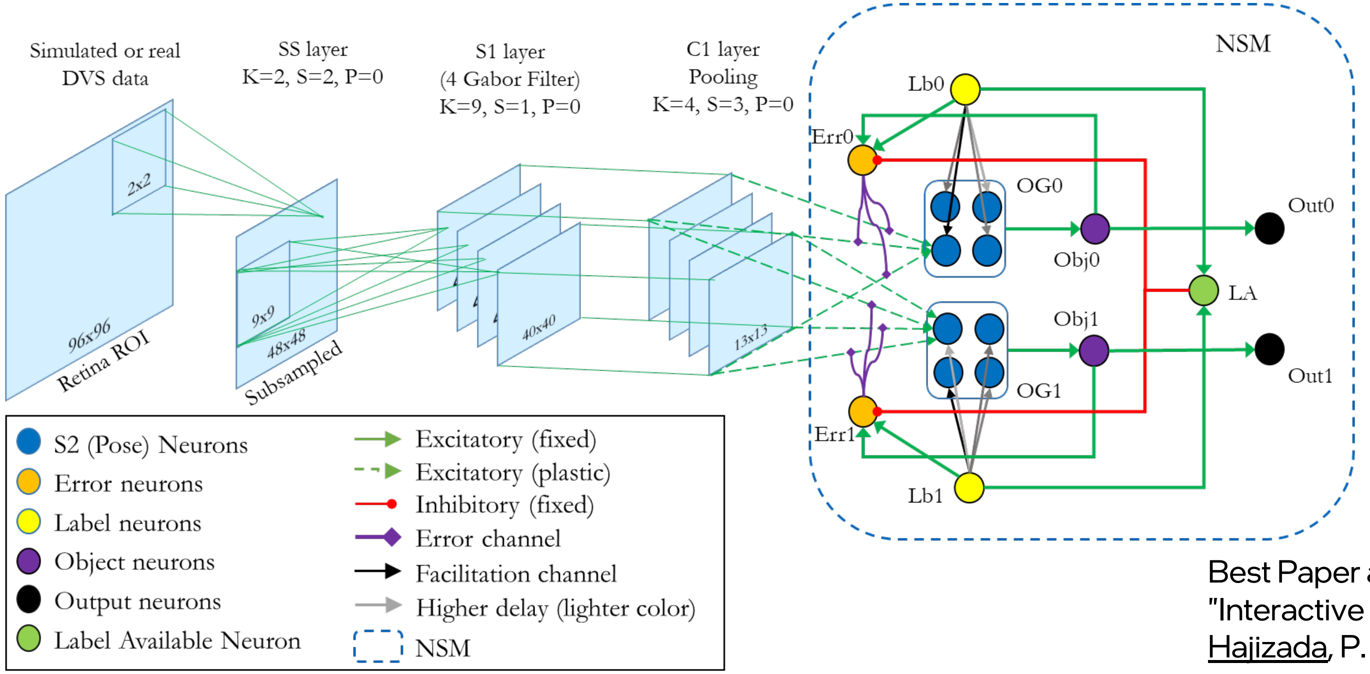
3) A.J. Ijspeert, Central pattern generators for locomotion control in animals and robots: A review. Neural Networks, vol. 21/4, pp. 642-653, 2008

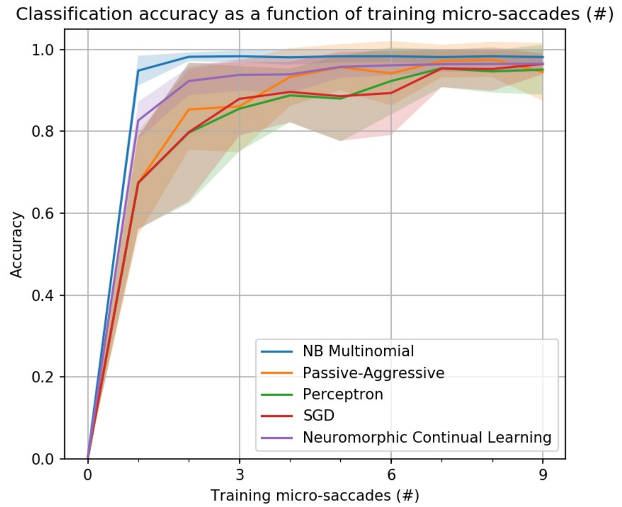
Example: object learning

- Learning objects in a natural way







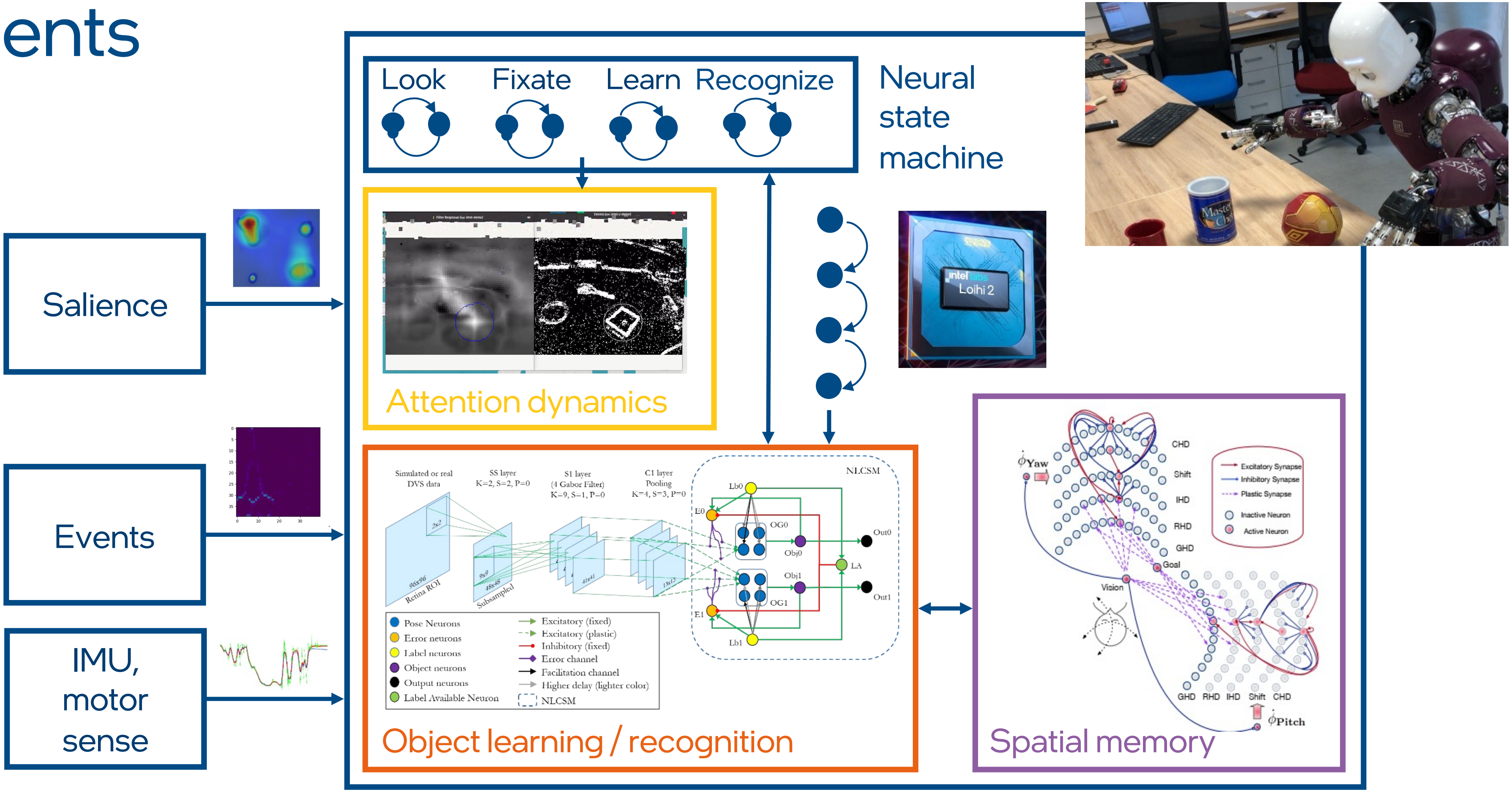


- 200x better energy per learning instance and up to 150x for inference
- The best execution time for learning an instance and being on par with other methods in inference time

Best Paper at International Conference on Neuromorphic Systems (ICONS): "Interactive continual learning for robots: a neuromorphic approach," E. Hajizada, P. Berggold, M. Iacono, A. Glover, Y. Sandamirskaya

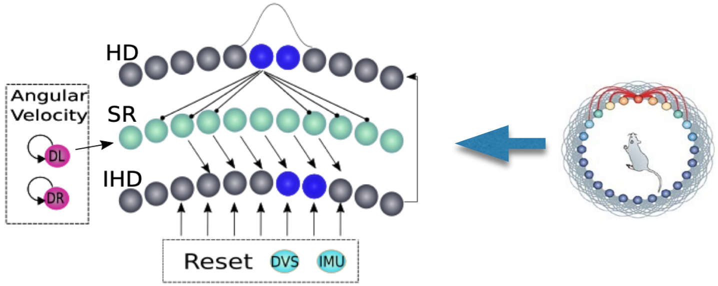
Notice: All experiments run on a machine with Intel(R) Core(TM) i7-9700K CPU @ 3.60GHz x8, 32GB RAM running Ubuntu 20.04.2 LTS, python 3.8.10, and v. 1.0.0 of the Intel NxSDK in Loihi 1 hardware. All performance measurements are based on testing as of October 2021 and may not reflect all publicly available security updates. Results may vary.

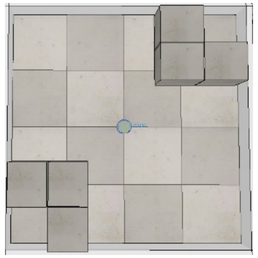
Combining with other behavioral elements

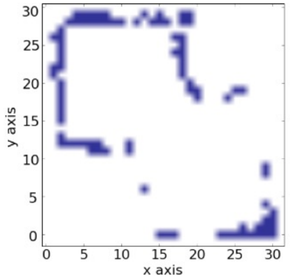


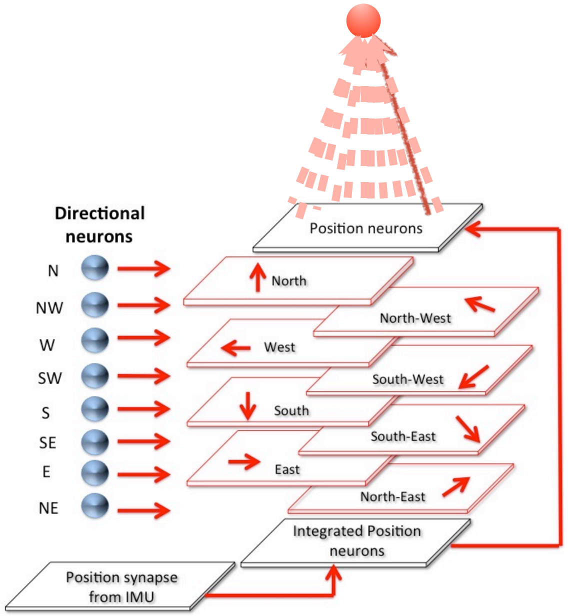
Spatial memories: forming and correcting a memory

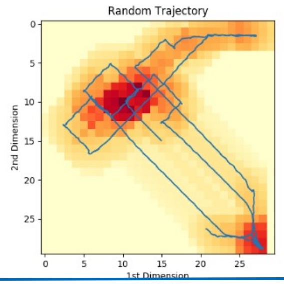
Place cells, Grid cells

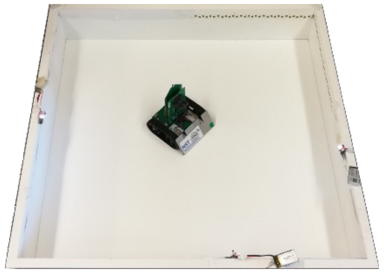


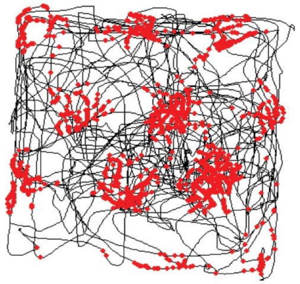






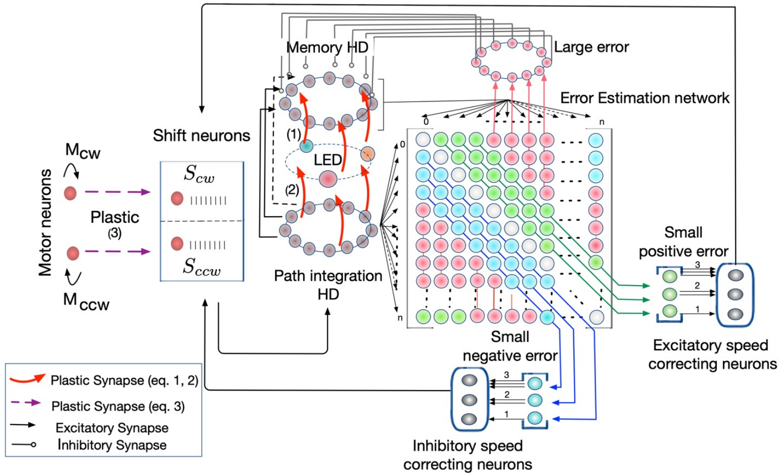


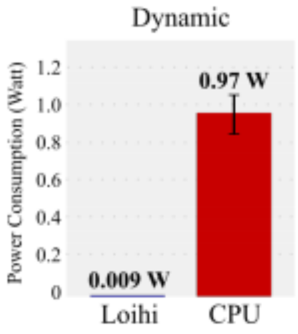




Moser et al. Annual Review of Neuroscience 2008

Error monitoring and correction



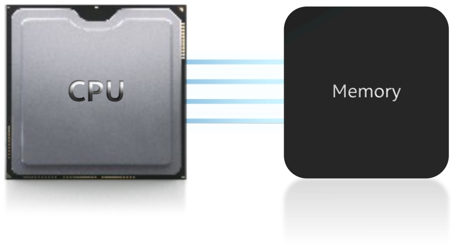


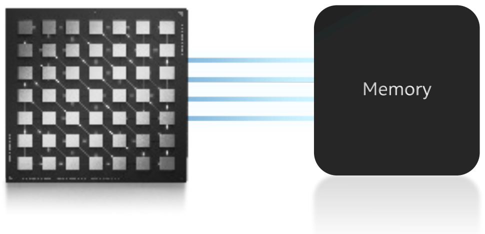
X100 more energy efficient compared to Gmapping on CPU (i7-4850HQ) on 1D SLAM

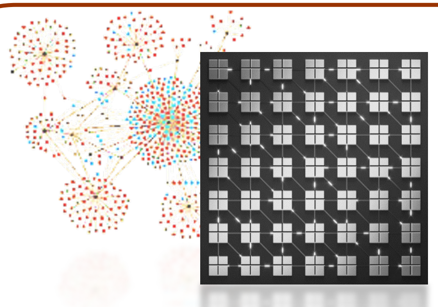
Kreiser et al, ISCAS 2018; Kreiser et al, IROS 2018, 2019; Kreiser at al, RAINR 2019;

Tang, Michmizos, ACM Proc., 2018

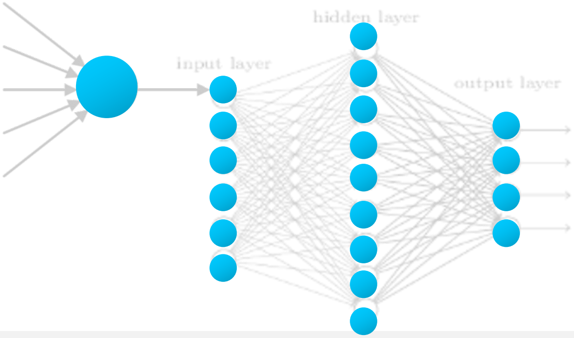
Implementing neural architectures efficiently

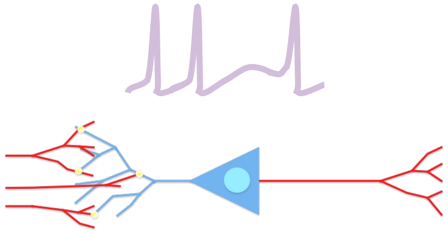






SEQUENTIAL THREADS OF CONTROL	PARALLEL DENSE COMPUTE	PARALLEL SPARSE COMPUTE
SYNCHRONOUS CLOCKING	SYNCHRONOUS CLOCKING	ASYNCHRONOUS EVENT-BASED SPIKES
PROGRAMMING BY ENCODING ALGORITHMS	OFFLINE TRAINING USING LABELED DATASETS	LEARN ON THE FLY THROUGH LOCAL LEARNING RULES





Neuromorphic hardware marketplace

Digital Chips

TrueNorth (IBM)

Loihi (Intel)

Tianjic (Tsinghua U)

Darwin 2 (Zhejian Lab)

SpiNNaker 2 (Human Brain Project)

DynapCNN (SynSense / Baidu)

GrAI One (GrAI Matter Labs)

Akida (BrainChip)

Mixed analog-digital Chips

ROLLS (ETH Zurich)

Braindrop (Stanford)

DynapSEL (SynSense / Baidu)

Innatera (Delft)

Large-Scale Systems

4M-neuron **BrainScaleS** (HBP)

64M-neuron **NS16e-4** (IBM)

1M-core **SpiNNaker** (HBP)

100M-neuron **Pohoiki Beach/Springs** (Intel)

120M-neuron **Darwin Mouse** (Zhejiang U)

Event-based Vision

DAVIS240C (IniVation) 18um

CeleX-IV (CelePixel) 18um

DVS-Gen2 (Samsung) 9um

DVS-Gen3 (Samsung) 6um

Gen4 (Prophesee/Sony)

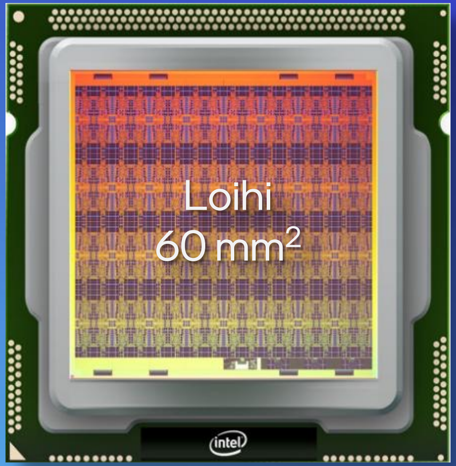
FENCE (DARPA) event-based IR

2014 2015 2016 2017 2018 2019 2020 2021

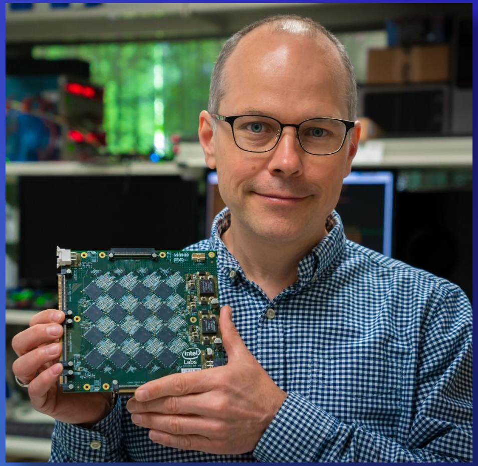
Five years ago, Intel Labs announced the Loihi neuromorphic test chip

Our mission: Pioneer a new programmable computing technology inspired by a modern understanding of the brain

Loihi Neuromorphic Research Chip



Nahuku board with 32 Loihi chips

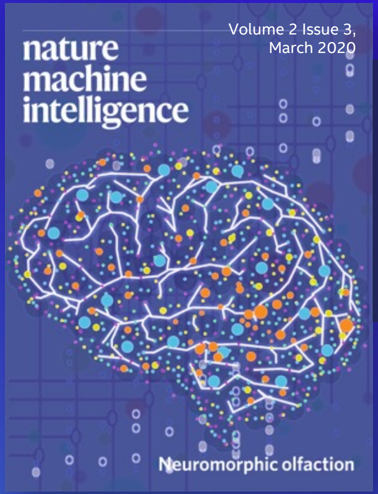


Pohoiki Springs system with 768 Loihi chips

Research Community with 180+ members

Image: intel.com/content/www/us/en/research/neuromorphic-community.html

Loihi application proof points



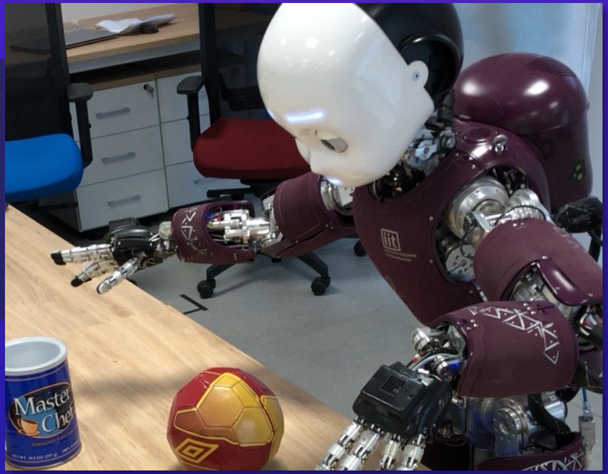
Olfaction-inspired odor recognition and learning

3000x more data efficient learning than a deep autoencoder

Scene understanding

Integrated behaviors: Object recognition, tracking, learning

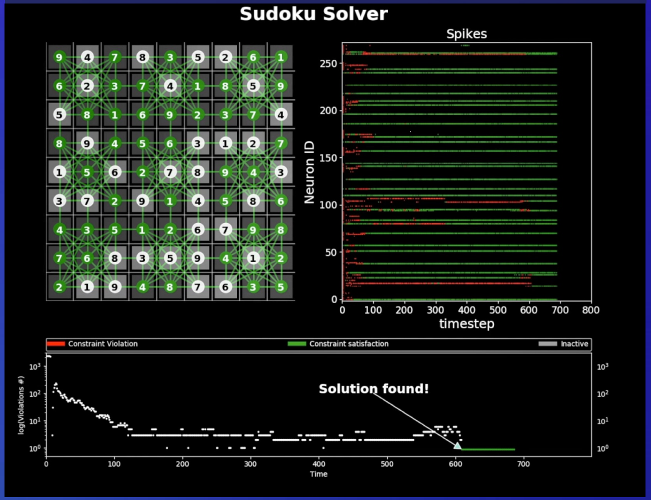
100x lower power SLAM vs CPU



Combinatorial optimization

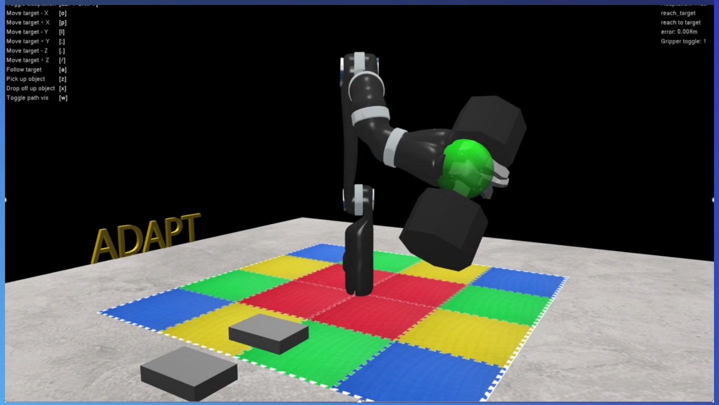
(CSP, SAT, ILP, QP)

2,800x lower energy and 44x faster vs CPU



Adaptive robotic arm control

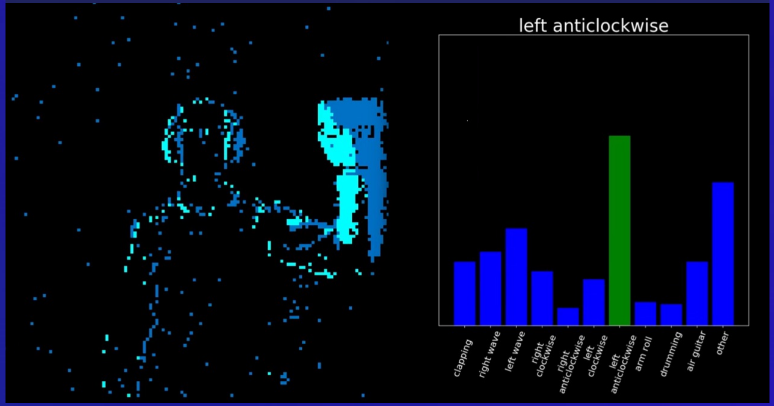
40x lower power, 50% faster vs GPU



Gesture recognition + learning

Loihi + DAVIS 240C camera

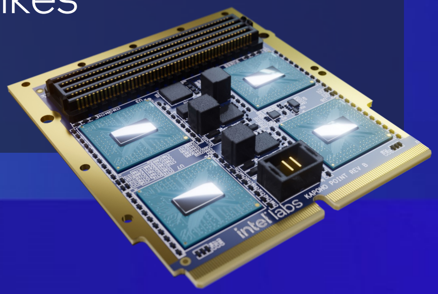
60 mW total power, 15 mW dynamic



Last year, Intel entered a new era with Intel Loihi 2 and open-source Lava framework

* specs and configuration details can be found at intel.com/neuromorphic

- Up to 10x faster processing capability*
- Up to 60x more inter-chip bandwidth*
- Up to 1 million neurons with 15x greater resource density*
- 3D scalable
- Native ethernet
- Programmable neurons
- Graded spikes



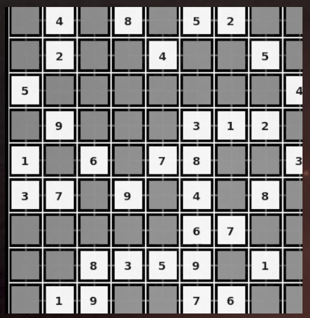
LAVA

- Event-based communication
- Multi-Paradigm
- Multi-Abstraction
- Multi-Platform

Open-Source and Community-Driven

Multi-Paradigm

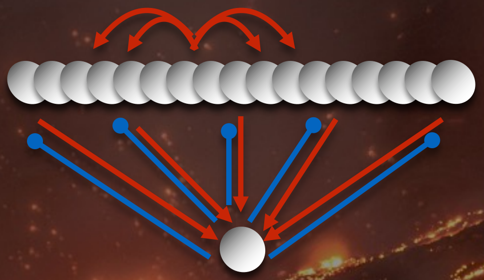
Optimization



LCA, Stochastic SNNs
LASSO, QP,
CSP, ILP, QUBO

+ model learning

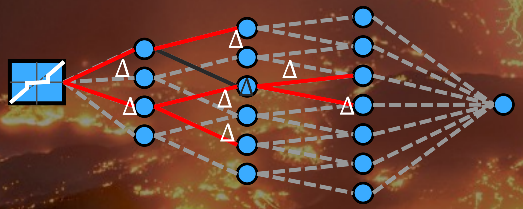
Neural Attractors



Dynamic Neural Fields,
Continuous Attractor NNs,
WTA

+ associative learning

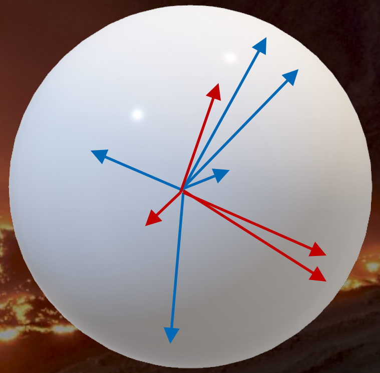
Deep Learning



ANN->SNN rate-coded conversion,
Directly trained SNN ConvNets
Sigma-Delta Neural Networks
TTFS- and Phase-coded SNNs

+ gradient learning

Vector Symbolic



HRRs, MAPs,
Binary Spatter Codes,
Sparse Block Codes,
Resonator Networks

+ HD learning

Many others to come: NEF, Reservoir Computing, STICK, Equilibrium Propagation, evolutionary, ...

Latest Lava Milestones and Results

- Intel added support for Loihi 2 features including **programmable neurons, graded spikes, and continual learning**.
- With the latest release of Lava (v0.5) and Kapoho Point, Intel Labs achieved **15x improved energy efficiency and up to 12x faster throughput** for a deep learning application.

	Loihi 1 SNN [3]	Loihi 2 SNN [2]		Loihi 2 SDNN [1]	
Mean-Square-Error	0.049	0.049	1	0.037	32% lower
Neuron cores	368	70	5x smaller	70	5x smaller
Latency (ms)	15.5	2.56	6x faster	1.22	9-12x faster
Throughput (fps)	808	4877		7404	
Energy (uJ/frame)	1770	270	6.5x better	120	15x more efficient
TOPS/W (DNN equiv)	0.02	0.13		0.28	

Results may vary.

[1] Loihi 2 SDNN results based on Lava v0.5 benchmarks in September, 2022 of 9-layer PilotNet DNN inference workload implemented as a sigma-delta neural network on Loihi 2. Equivalent DNN op counts calculated from a conventional DNN implementation with the same topology and same number of 8-bit parameters. See Bojarski, Mariusz et al. "End to end learning for self-driving cars." arXiv preprint arXiv:1604.07316 (2016).

[2] Loihi 2 SNN measurements were obtained on Oheo Gulch board ncl-og-06 using an internal version of NxSDK.
[3] Loihi 1 SNN measurements were obtained on Nahuku 32 board ncl-ghrd-01 using NxSDK v1.0.0

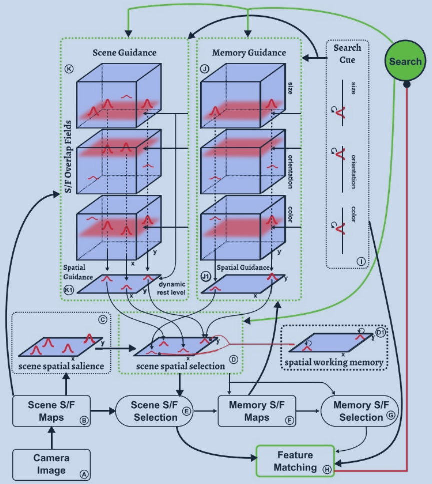
Redefining Artificial Intelligence with Neuromorphic Computing

Diversity of neuronal "algorithms"

Symbolic processing

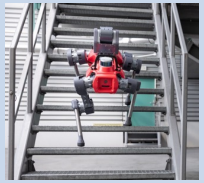
- Representation: Sensing, memory
- Evaluation of options: Optimization, planning
- Decision making
- Action: active sensing, sensing for acting

Artificial embodied intelligence



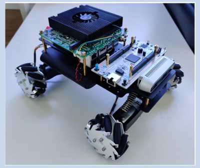
- Building intelligent neural architectures
- With building blocks inspired by bio-computing principles

locomotion



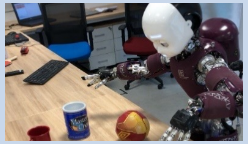
drones

vehicles



- Enabling smart autonomous systems

arms, grippers



cognitive, humanoids

autonomous spaces

Intel Neuromorphic Research Community

To join:
inrc_interest@intel.com

Collaborating to Accelerate the Research

INRC includes over ~~120~~ groups
170

Legal Information

Performance varies by use, configuration and other factors. Learn more at www.Intel.com/PerformanceIndex.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Your costs and results may vary.

Results have been estimated or simulated.

Intel technologies may require enabled hardware, software or service activation.

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

References and System Test Configuration Details

[Task 1] P Blouw et al, 2018. arXiv:1812.01739

[Task 2] TY Liu et al, 2020, arXiv:2008.01380

[Task 3] KP Patel et al, "A spiking neural network for image segmentation," *submitted, in review*, Aug 2020.

[Task 4] **Loihi:** Nahuku system running NxSDK 0.95. CIFAR-10 image recognition network trained using the SNN-Toolbox (code available at https://snntoolbox.readthedocs.io/en/latest). **CPU**: Core i7-9700K with 32GB RAM, **GPU**: Nvidia RTX 2070 with 8GB RAM. OS: Ubuntu 16.04.6 LTS, Python: 3.5.5, TensorFlow: 1.13.1. Performance results are based on testing as of July 2020 and may not reflect all publicly available security updates.

[Task 5] **Loihi**: Nahuku system running NxSDK 0.95. Gesture recognition network trained using the SLAYER tool (code available at https://github.com/bamsumit/slayerPytorch). Performance results are based on testing as of July 2020 and may not reflect all publicly available security updates. **TrueNorth:** Results and DVS Gesture dataset from A. Amir et al, "A low power, fully event-based gesture recognition system," in IEEE Conf. Comput. Vis. Pattern Recog. (CVPR), 2017.

[Task 6] T. Taunyazov et al, 2020. RSS 2020

[Task 7] Bellec et al, 2018. arXiv:1803.09574. **Loihi**: Loihi: Wolf Mountain system running NxSDK 0.85. **CPU**: Intel Core i5-7440HQ, with 16GB running Windows 10 (build 18362), Python: 3.6.7, TensorFlow: 1.14.1. **GPU**: Nvidia Telsa P100 with 16GB RAM. Performance results are based on testing as of December 2018 and may not reflect all publicly available security updates.

[Task 8] T. DeWolf et al, "Nengo and Low-Power AI Hardware for Robust, Embedded Neurorobotics," Front. in Neurorobotics, 2020.

[Task 9] Loihi Lasso solver based on PTP Tang et al, "Sparse coding by spiking neural networks: convergence theory and computational results," arXiv:1705.05475, 2017. **Loihi**: Wolf Mountain system running NxSDK 0.75. **CPU**: Intel Core i7-4790 3.6GHz w/ 32GB RAM running Ubuntu 16.04 with HyperThreading disabled, SPAMS solver for FISTA, http://spams-devel.gforge.inria.fr/.

[Task 10] G Tang et al, 2019. arXiv:1903.02504

[Task 11] EP Frady et al, 2020. arXiv:2004.12691

[Task 12] Loihi graph search algorithm based on *Ponulak F., Hopfield J.J. Rapid, parallel path planning by propagating wavefronts of spiking neural activity. Front. Comput. Neurosci. 2013.* **Loihi**: Nahuku and Pohoiki Springs systems running NxSDK 0.97. **CPU**: Intel Xeon Gold with 384GB RAM, running SLES11, evaluated with Python 3.6.3, NetworkX library augmented with an optimized graph search implementation based on Dial's algorithm. See also http://rpg.ifi.uzh.ch/docs/CVPR19workshop/CVPRW19_Mike_Davies.pdf

[Task 13] **Loihi**: constraint solver algorithm based on *G.A. Fonseca Guerra and S.B. Furber, Using Stochastic Spiking Neural Networks on SpiNNaker to Solve Constraint Satisfaction Problems. Front. Neurosci. 2017.* Tested on the Nahuku 32-chip system running NxSDK 0.98. **CPU**: Core i7-9700K with 32GB RAM running Coin-or Branch and Cut (https://github.com/coin-or/Cbc). Performance results are based on testing as of July 2020 and may not reflect all publicly available security updates.

Loihi 2 Performance Analysis Details

[2] Based on comparisons between barrier synchronization time, synaptic update time, neuron update time, and neuron spike times between Loihi 1 and 2. Loihi 1 parameters measured from silicon characterization (see below); Loihi 2 parameters measured from both silicon characterization with N3B1 revision and pre-silicon circuit simulations using back-annotated timing for Loihi 2.

[3] Based on Lava simulations in September, 2021 of a nine-layer variant of the PilotNet DNN inference workload implemented as a sigma-delta neural network on Loihi 2 compared to the same network implemented with SNN rate-coding on Loihi. The Loihi 2 SDNN implementation gives better accuracy than the Loihi 1 rate-coded implementation.

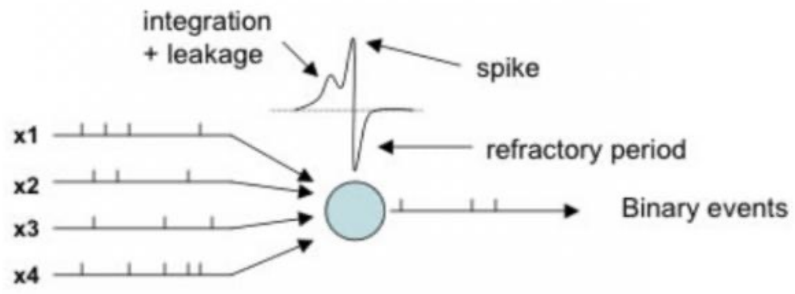
[4] Circuit simulations of Loihi 2's wave pipelined signaling circuits show 800 Mtransfers/s compared to Loihi 1's measured performance of 185 Mtransfers/s.

[5] Based on analysis of 3-chip and 7-chip Locally Competitive Algorithm examples.

The Lava performance model for both chips is based on silicon characterization in September 2021 using the Nx SDK release 1.0.0 with an Intel Xeon E5-2699 v3 CPU @ 2.30 GHz, 32GB RAM, as the host running Ubuntu version 20.04.2. Loihi results use Nahuku-32 system ncl-ghrd-04. Loihi 2 results use Oheo Gulch system ncl-og-04. Results may vary.

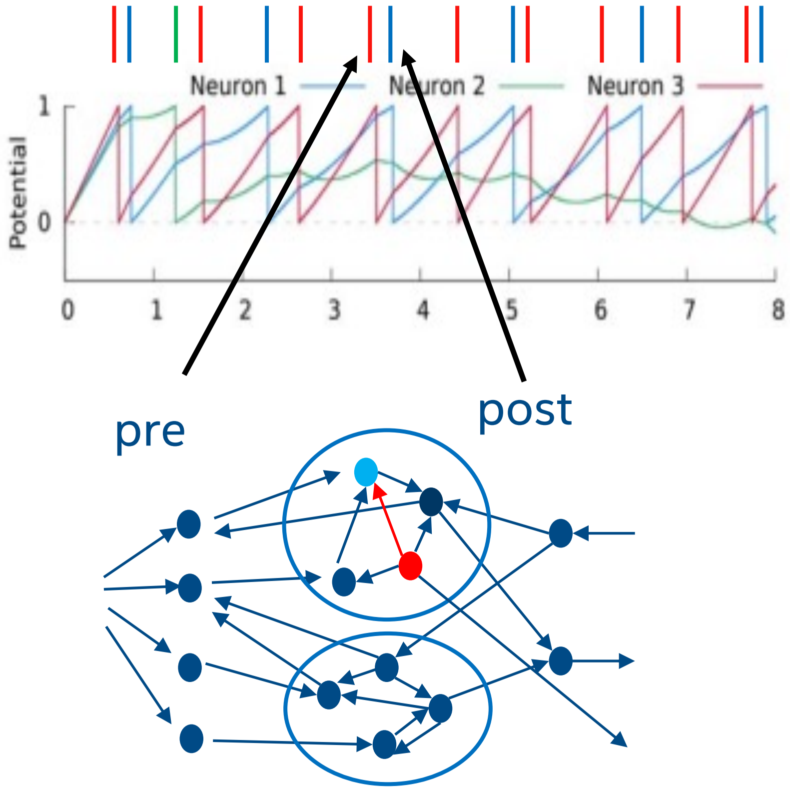
Neuromorphic computing: Core elements

Spiking neuron: leaky integrate and fire

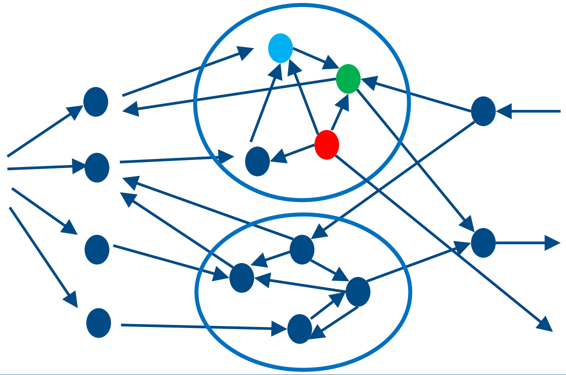


- Time is explicitly included in computation
- Events (spikes) transmit activation
- Spatial-temporal patterns

Learning: synaptic plasticity



- Local learning rules



Network topology

- Fine-grained parallelism
- Modularity, recurrence