EECS 507: Introduction to Embedded Systems Research
Midterm Exam
15 Dec 2022

Robert Dick

- You have 110 minutes for the exam.
- Neither books, notes, nor internet may be used.
- Calculators are permitted, so long as are used only for calculation, not for storing notes.
- There are answer length limits to control exam duration. For the sake of fairness, if you exceed the length limits, I will evaluate only the portion of the answer within the length limit.
- No communicating with anybody except the teacher about the exam. This remains true even after you have submitted the exam. I'll tell you via Piazza when it is O.K. to discuss. This is to prevent problems if a student needs to take the exam late due to illness, injury, or exam time conflict.
- Skim all the questions before starting so you can budget your time. They have different difficulties, but each is worth the same credit; that's intentional.
- $0\,°\mathrm{C} = 273\,\mathrm{K}$.

Printed name

Sign below to acknowledge the Engineering Honor Code: "I have neither given nor received aid on this examination, nor have I concealed a violation of the Honor Code."

Signature

1. **Breadth questions on reading assignments**

10 (a) [P. Kanerva, "Hyperdimensional computing: an introduction to computing in distributed representation with high-dimensional random vectors," *Cognitive Computing*, vol. 1, Jan. 2009]
What property is a main enabler of hyperdimensional computing?
- ◯ Exponentially increasing bit weights allow compact representations of scalars such as activations.
- ◯ High-dimensionality random vectors can be added using a standard two's-complement adder.
- ◯ Hyperdimensional computing produces canonical representations of objects, e.g., a number, a color, or an animal, that are consistent across hyperdimensional computing implementations.
- ◯ Hyperdimensional computing emulates the efficient filtering carried out within human retinas.
- ◯ A randomly selected pair of high-dimensionality random vectors is almost always almost orthogonal.

10 (b) [A. Bonde, J. R. Codling, K. Naruethep, Y. Dong, W. Siripaktanakon, S. Ariyadech, A. Sangpetch, O. Sangpetch, S. Pan, H. Y. Noh, and P. Zhang, "PigNet: failure-tolerant pig activity monitoring system using structural vibration," in *Proc. Int. Symp. Information Processing in Sensor Networks*, May 2001]
What was the primary purpose of PigNet?
- ◯ Tracking the locations of pigs over time.
- ◯ Resisting damage from water and pig waste.
- ◯ Evaluating the energy efficiency implications of using geophones for vibration sensing.
- ◯ Detecting the times when pig nursing and lying occur.
- ◯ Thailand.

10 (c) [M. Chen, D. Gündüz, K. Huang, W. Saad, M. Bennis, A. V, Feljan, and H. V. Poor, "Distributed learning in wireless networks: recent progress and future challenges," *IEEE J. on Selected Areas in Communication*, vol. 39, no. 12, Dec. 2021]
Which one of the following is true of federated learning, using the conventional definition in the paper?
- ◯ Each device maintains a distinct machine learning model tuned to its local environment.
- ◯ The speed and accuracy of learning is independent of wireless communication conditions.
- ◯ The data distributions for the clients may differ.
- ◯ It guarantees that privacy is preserved.
- ◯ Each device is associated with a different server.

10 (d) R. Banakar, S. Steinke, B.-S. Lee, M. Balakrishnan, and P. Marwedel, "Scratchpad memory: A design alternative for cache on-chip memory in embedded systems," in *Proc. Int. Wkshp. Hardware/Software Co-Design*, May 2002, pp. 73–78
What statement is true regarding the use of scratchpad memory when implementing a multi-layer perceptron using an instruction processor?
- ◯ The regular computational structure, with limited conditionals, is well suited to compile-time memory access pattern analysis.
- ◯ Scratchpad memory would be more energy-efficient than cache because it can be implemented on the same die as the instruction processor.
- ◯ The scratchpad memory would require less die area because it needs no decoder.
- ◯ Scratchpad memory is typically implemented using area-efficient, deep-trench capacitors.
- ◯ This application has no hard deadlines and therefore scratchpad memory would bring no benefit.

$\boxed{10}$ (e) Y. Zhu, A. Samajdar, M. Mattina, and P. Whatmough, "Euphrates: Algorithm-SoC co-design for low-power mobile continuous vision," arXiv, Tech. Rep., Apr. 2018
Which one of the following statements is false, regarding Euphrates?

- ◯ A region of interest (ROI) is divided into sub-ROIs to accommodate non-rigid or rotating objects.
- ◯ The motion controller coordinates the CNN engine without interrupting the CPU.
- ◯ The CNN accelerator is based on a systolic array.
- ◯ A custom accelerator is used to compute motion vectors.
- ◯ Euphrates reduces energy consumption by more than two orders of magnitude compared to conventional continuous vision pipeline.

2. Student projects

$\boxed{10}$ (a) Which statement is false regarding the FPGA-based hardware-in-the-loop testing project?

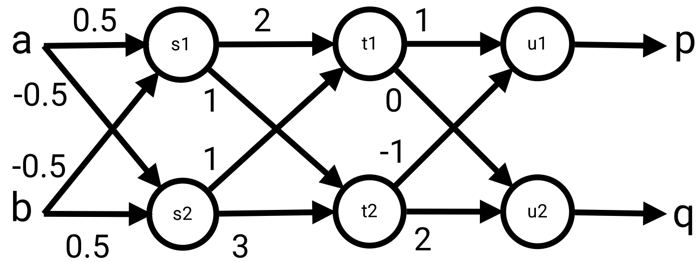
- ◯ The framework emulates the embedded system's operating environment.
- ◯ The entire embedded system and its operating environment are simulated to allow firmware and application testing and validation.
- ◯ The FPGA configuration will depend on the hardware being emulated.
- ◯ One of the test cases involved robot maze navigation.
- ◯ Open-source synthesis tools were used to produce FPGA programming bitstreams.

$\boxed{10}$ (b) In the electronic nose student project evaluation, why was the accuracy for one type of food higher than for the other type?

- ◯ The food with lower accuracy did not generate carbon dioxide.
- ◯ The food with lower accuracy decomposed too slowly to observe any change during the testing period.
- ◯ The food with higher accuracy was dead.
- ◯ Only one of the foods generated enough H_2S to be detected.
- ◯ Only one food caused enough movement to be detected by the accelerometer.

3. Lecture-related questions

10 (a) The following multi-layer neural network uses linear activation functions.



Its inputs a and b are in range [0, 1]. Write the mathematical functions associated with output p

and output q.

$\boxed{10}$ (b) Consider the design of an IoT edge device that must classify the genus of a flying insect based on data from a microphone array. There are three alternative implementation approaches

1. Do classification entirely on the edge device and transmit (using negligible energy) a compact classification code to the cloud. In this case, classification occupies the processor for 20 s.
2. Send raw data to the cloud. In this case, 10,000 kb of data must be transmitted.
3. Losslessly compress the audio data and transmit it to a cloud server for learning and inference. In this case, compression occupies the processor for 1 s and the compression ratio is 0.1.
4. Do feature extraction on the edge device and transmit the feature vector to the cloud for learning and inference. In this case, 200 kb of data must be transmitted and feature extraction occupies the processor for 5 s.

Each approach achieves similar accuracy. Regardless of design approach, sampling requires that the processor spend 0.1 s active. The LPWAN technology in use supports a data rate of 20 kb/s and has a transmission-mode power consumption of 10 mW. Whenever transmitting, the processor must also be active. The goal is to minimize energy consumption on the edge device. Each sampling cycle takes 1 minute.

The following table gives relevant power consumptions.

Component	Active power (mW)	Sleep power (mW)
Processor	1,000	300
Transmitter	10	0

Which design approach is optimal?

- ◯ Do everything on the edge device.
- ◯ Send raw data to the cloud.
- ◯ Compress and classify in the cloud.
- ◯ Send only the features and classify in the cloud.

What is the energy consumption per sensing cycle using the optimal approach?

10 (c) You are designing an embedded system that will be subject to pin monitoring attacks, in which an attacker may use an FPGA or logic analyzer to monitor the voltages of all printed circuit board traces. In your design, there are certain secret intermediate values that will be computed during inference and that must never be leaked on the printed circuit board traces. The embedded system will have DRAM on a separate chip from the microprocessor. You can choose either a processor with on-chip scratchpad memory or a processor with on-chip cache. The processor core is identical for these two cases.

Which would you choose?

- ◯ Scratchpad memory.
- ◯ Cache.

Using at most one sentence or phrase, explain why.

10 (d) What is the name of the algorithm that linearly projects vectors into a new space of the same dimensionality as the original space with orthonormal dimensions ordered by decreasing eigenvalues of the input data matrix?

- ◯ Radial basis function.
- ◯ Principle component analysis.
- ◯ K-means.
- ◯ $A^\star$ search.
- ◯ Spectral analysis.

10 (e) In the human brain, when dendrites proximal to an axon fire in synchrony, how does this most commonly effect the timing of the neuron's next pulse?

- ◯ Accelerates.
- ◯ Delays.

10 (f) What statement about autoencoders is false?

- ◯ They typically have narrower layers approximately midway between inputs and outputs.
- ◯ When well trained, their outputs estimate or reconstruct their inputs.
- ◯ They must have at least four layers.
- ◯ They can be used to learn distribution-appropriate latent-space lossy data compression from a data set.
- ◯ They do not contain recurrent units.

10 (g) What two types of CMOS power consumption are generally most relevant to embedded systems designers?

- ◯ GIDL leakage.
- ◯ Switching.
- ◯ Subthreshold leakage.
- ◯ Short-circuit.
- ◯ Gate leakage.
- ◯ Junction leakage.

$\boxed{10}$ (h) Name the phenomena resulting from the fact that power consumption is a superlinear function of temperature?

- ◯ Asynchronous element finite difference techniques.
- ◯ dI/dt effect.
- ◯ Runge-Kutta decomposition.
- ◯ Newton-Raphson iteration.
- ◯ Thermal runaway.