Introduction to Embedded Systems Research
Final Exam

Robert Dick

30 April 2019

Closed book. Closed notes. No calculators or other computers.
If you write lightly with pencil, I may not see your answers or work.

Name: Robert Dick

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."

Using at most one sentence, each, describe the main contribution of each of these five research papers.

[5] 1. 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.

Used motion vectors to determine which regions of the scene were superfluous, eliminating energy waste doing inference on those regions. (Also made other contributions, some of which received full marks.)

[5] 2. E. Ronen, A. Shamir, A.-O. Weingarten, and C. O'Flynn, "IoT goes nuclear: Creating a ZigBee chain reaction," in *Proc. Symp. on Security and Privacy*, May 2017.

Exploited bug in implementation of ZigBee Touchlink protocol and extracted a global AES key to enable arbitrary firmware updates of Philips Hue bulbs and prove that a worm could spread.

[5] 3. P. M. Sheridan, F. Cai, C. Du, W. Ma, Z. Zhang, and W. D. Lu, "Sparse coding with memristor networks," *Nature Nanotechnology*, vol. 12, Aug. 2017.

They demonstrated the use of a memristor crossbar to implement a sparse coding algorithm for images, in which lateral inhibition was an important mechanism.

[5] 4. P. Coussy, C. Chavet, H. Wouafo, and L. Conde-Canecia, "Fully binary neural network model and optimized hardware architectures for associative memories," *ACM J. on Emerging Technologies in Computing Systems*, vol. 11, no. 4, Apr. 2015.

They demonstrated a fully binary model that eliminated the need for inefficient WTA voting and also reduced storage requirements by reducing redundancy in weight storage.

[5] 5. L. Zhang, B. Tiwana, Z. Qian, Z. Wang, R. P. Dick, Z. M. Mao, and L. Yang, "Accurate online power estimation and automatic battery behavior based power model generation for smartphones," in *Proc. Int. Conf. Hardware/Software Codesign and System Synthesis*, Oct. 2010, pp. 105–114.

Showed a method of generating EB power models using only built-in battery voltage sensors and knowledge of battery discharge curve.

[5] 6. For the student project on using parallel recombinative simulated annealing to design printed circuit board trace antennas, a single optimization run took many hours. Why? In other words, what was the majority of time spent on?

- ○ Calculating the results of Boltzmann trials.
- ○ Random number generation.
- ● Running an electromagnetics simulator.
- ○ Calculating interactions with printed circuit board traces used for power distribution.

[5] 7. Which of the following errors can be detected via model checking?

- ● Failure of the design to always meet the requirements in the specifications.
- ○ Hardware component faults resulting from manufacturing process variation.
- ○ Mismatches between specification and designer intentions.

[10] 8. List three reasons (a few words each, at most) that high power consumption can cause problems in embedded systems. Helpful note: Rephrasing the same reason doesn't turn it into different reasons.

1) Leads to high temperature, which accelerates wear.
2) Reduces battery lifespan.
3) Either produces faults due to IR drop or requires more carefully designed/more expensive power delivery systems. ↖ I took much shorter answers. Trying to give detail.

[5] 9. Consider the following table.

Wireless Communication Technologies

Technology	Power (mW)	Range (m)	Typical rate (kb/s)
4G	1,000	70,000	10,000
5G	1,000	40,000	100,000
WiFi / 802.11(g)	250	140	20,000
Zigbee / 802.15.4	50	10–1,500	100
LoRaWAN	10	15,000	20
NB-IoT	100	15,000	250

Which communication technology is most energy efficient per bit, and what is its energy cost per bit? Show your work.

5G. 10nJ/bit.

$$\frac{1W}{100Mb/s} = 10nJ/b$$

[5] 10. List four methods of improving the energy efficiencies of deep neural networks, using at most a few words each.

1) Prune edges/weights.
2) Reduce weight precision.
3) Compress weights, e.g., using indices into tables.
4) Application-specific hardware supporting parallel MACs, e.g., FPGAs or GPUs.

[5] 11. In a convolutional neural network, how many multiply-accumulate operations are required to apply a 5×5 kernel to a 100×100 two-dimensional image? You may neglect computations related to indexing operations.

For valid convolution, output is 96×96. Assume stride = 1.

$96^2 \cdot 5^2$

$$\begin{array}{r} 96 \\ \times\ 96 \\ \hline 576 \\ +\ 8640 \\ \hline 9216 \end{array} \qquad \begin{array}{r} 9216 \\ \times\ 25 \\ \hline 46080 \\ 184320 \\ \hline 230{,}400 \end{array}$$

230,400

Also accepted answers for "full" and "same" convolution.

[5] 12. In a convolutional neural network, how many multiply-accumulate operations are required to evaluate a single neuron in the second of two fully-connected layers, each containing 1,000 neurons? You may neglect computation related to applying the activation function to the sum.

1,000. 1,001 with the explicit statement that a bias term was included would also be accepted.