Midterm Exam for Computer Networks (CNT 6215)

Fall 2014

>>> SOLUTIONS <<<

Welcome to the Midterm Exam for *Computer Networks*. Read each problem carefully. There are eight required problems (each worth 12 points – you get 4 points for correctly following these instructions). There is also an additional extra credit question worth 10 points. You may have with you a calculator, pencils and/or pens, erasers, blank paper, and one 8.5 x 11 inch "formula sheet". On this formula sheet you may have anything you want (definitions, formulas, homework answers, old exam answers, etc.) as **handwritten by you in pencil or ink** on both sides of the sheet. Photocopies, scans, or computer generated and/or printed text are not allowed on this sheet. Note to tablet (iPad, etc.) users – you may **not** print-out your handwritten text for the formula sheet. You have 75 minutes for this exam. **Please use a separate sheet of paper for the answer to each question**. Good luck and be sure to show your work!

Problem #1

Each sub-problem worth 3 points.

Answer the following questions regarding the basics principles and concepts of networks.

a) Sketch the 7-layer OSI reference model. In one sentence, or less, describe the function of (or service provided by) each layer.

```
   +--------------+
 7 | Application  | --  Provides access to user applications
   +--------------+
 6 | Presentation | --  Provides data independence
   +--------------+
 5 |   Session    | --  Manages end-to-end connections
   +--------------+
 4 |  Transport   | --  Provides reliable end-to-end data transport
   +--------------+
 3 |   Network    | --  Maintains point-to-point connections
   +--------------+
 2 |   Data Link  | --  Provides reliable point-to-point data transport
   +--------------+
 1 |   Physical   | --  Transmission of bit stream
   +--------------+
```

b) Define *protocol* and *interface*. Give one example of each.

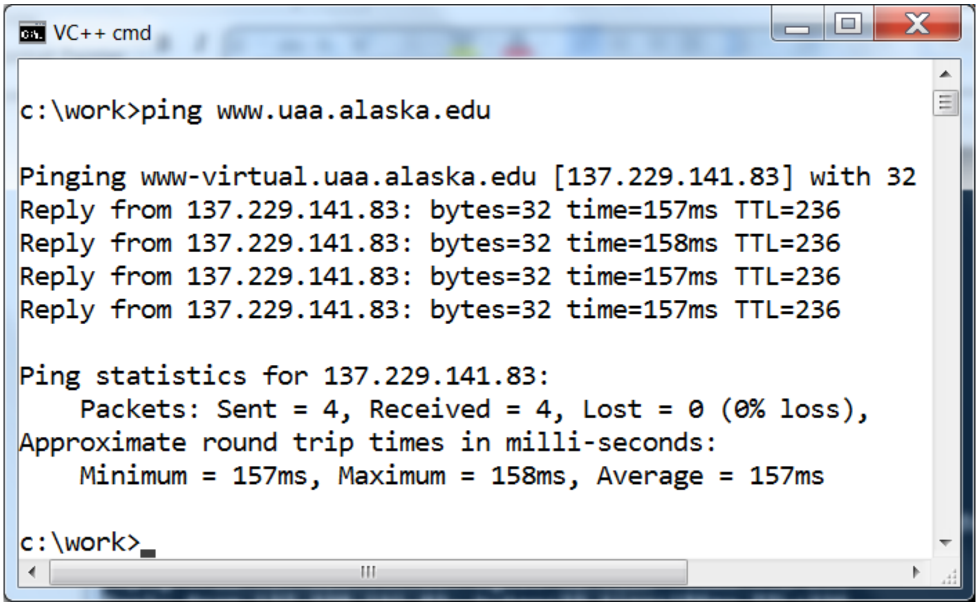
Protocol = Complete set of rules regarding information exchange between same level layers between sites. Example = HTTP.

Interface = Complete set of rules regarding information exchange between adjacent layers within a site. Example = Sockets.

c) What is the IETF and what is their mission?

The IETF is the Internet Engineering Task Force. "The mission is to make the Internet work better by producing high quality, relevant technical documents that influence the way people design, use, and manage the Internet."

d) Here is a screenshot showing a ping from USF to University of Alaska Anchorage. It is about 4000 miles to Anchorage from here. What percentage of the round-trip time in the below is due to propagation? What is the most likely primary cause of the remaining (non-propagation) delay? Carefully list any assumptions you make to answer this question.



Signal propagates at about the speed of light, which is about 1 nanosecond per foot or 5 microseconds per mile. At 4000 miles one-way the round-trip propagation delay is about 40 milliseconds. Assuming a fairly straight path from Tampa to Anchorage, about 120 millseconds of the shown round-trip delay is not due to propagation. It is most likely due to queueing delays in routers between Tampa and Anchorage.

Problem #2

Each sockets function is 1 pt.

Sockets is the standard API between Internet applications and the TCP (or UDP) protocol. Sockets is based on a client/server model. List the key sockets functions and for each function give a one sentence description of what the function does. The key functions are those that a working client/server sockets program could not be written without. Consider both streams and datagram service.

```
socket()- creates a socket of a specified type and allocates resources to it.
bind()- associates a socket with a server address
listen() – causes a bound socket to enter into a listening (for connections) state
connect()- used on client side, assigns a free local port to a socket
accept() – used on server side, accepts an incoming connection and creates a new socket
send() / recv() – send and receive data bytes for streaming (TCP)
sendto()/ recvfrom() – send and receive datagrams for datagram (UDP)
closesocket() – closes a socket and releases resources
```

Problem #3

U formula is 8 points, p formula is 2 points, calculation is 2 points

Consider a Stop-and-Wait protocol operating point-to-point across a single link with the following parameters:

- Message length = 50 bytes
- ACK length = 50 bytes
- Link rate = 100 kb/s
- Link length = 1000 miles
- Probability of a bit error = 10^{-3}
- The sender always has data (messages) to send
- Assume that ACKs are never lost.

What is the link utilization? Is the assumption of ACKs are never lost reasonable (explain your answer including why we make this assumption)?

We must note that t_{ack} is no longer negligible as it is the same value as t_{packet}. Thus, the link utilization will now be

$$U = \frac{(1-p)\cdot t_{fr}}{2\cdot t_{prop} + t_{packet} + t_{ack}}.$$

Where p is the probability of a packet loss and is determined as

$$p = 1 - (1 - \Pr[\text{bit error}])^{400}$$

(the 400 is the number of bits in a 50 byte message). Doing some calculations:

$$t_{packet} = t_{ack} = \frac{400}{100.0\cdot 10^3} = 4\ \text{ms}$$

$t_{prop} = 5\ \text{ms}$ (for 1000 miles at 5 microseconds per mile)

$$p = 1 - (1 - 10^{-3})^{400} = 0.3298$$

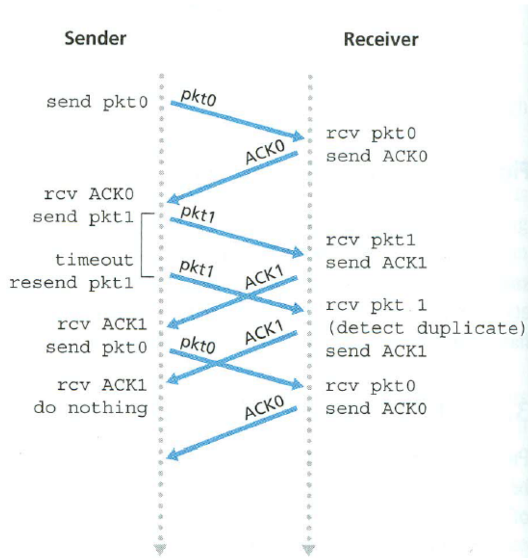
The above values plugged into the above U formula result in U = 14.9%.

The assumption on ACKs are never lost is not reasonable when the length of a message and ACK is the same – the probability of loss is thus the same also for each. We make this assumption for reasons of tractability – Deriving N_{retry} would be much more difficult if we allowed for the possibility of lost ACKs.

Problem #4

A basic SAW timing diagram is 6 points, the double ACK is 6 points

Consider the RDT protocol from Kurose and Ross (hint: the RDT protocol is a SAW protocol) discussed in my away lecture slides and covered in your reading. The keys RDT FSMs are in Appendix A. Sketch the timing diagram (that is, show the packet flows between the sender and receiver) for the case of a premature time-out. A premature time-out occurs when an ACK is "in flight" from receiver to sender and a time-out occurs in the sender before the ACK is received.



d. Premature timeout

From Kurose and Ross.

Problem #5 Each sub-problem is worth 4 points.

Answer the following questions regarding layer-2.

a) What are the four ways to share a channel?

```
Frequency division multiplexing (FDM), time division multiplexing (TDM), contention,
and token passing.
```

b) Describe the Ethernet CSMA/CD algorithm including the algorithm for determining the random delay for retransmission after a collision.

```
CSMA/CD:

(1) if (medium is idle) than xmit
(2) if (medium is busy) then wait until medium is idle and then xmit
(3) if (detect a collision) xmit a jam signal and then stop xmitting and then wait a
    random period of time (delay) then (1)

BEB:

while attempts < backoff_limit
  k := min(attempts, 10)
  r := rand(0, 2^k)
  delay := r * slot_time

slot_time is the minimum frame length (512 bits in IEEE 802.3).
```

c) What is the goal of the algorithm that determines the random delay in Ethernet?

```
The goal of BEB is to estimate, in a distributed fashion, how many nodes (Q) are
competing for the channel and to approximate Pr[xmit] = 1/Q as this is optimum (from
Metcalfe analysis).
```

Problem #6 Each sub-problem worth 3 points.

Answer the following questions regarding performance modeling and experimental design.

a) What is a model (precisely define it), what is the goal of building a model, and why do we build models (and not just study actual systems)?

```
"A model is a representation (physical, logical, or functional) that mimics another
object under study" (Molloy 1989). The goal of a building a model is to be able to
understand the system under study – ultimately, to be able to predict its performance.
We build models because they are cheaper, easier, faster, and/or safer to experiment
with compared to actual systems.
```

b) What are at least five performance measures of interest for computer networks? What is probably the most important (that is, the most interesting) performance measure?

```
Delay, Throughput, Utilization, Efficiency, Reliability, and Availability were noted in
class. Delay is probably the key performance measure of interest for computer networks.
```

c) What is the goal of output analysis as we discussed it in class (hint: think about sample mean and population mean)?

The goal is to be say something like “We are 95% confident that the true population mean lies between value1 and value2” where value1 and value2 are the sample mean plus/minus a confidence interval half width.

d) What is the advantage (and disadvantage) of a full-factorial experiment design compared to a simple experiment design?

Advantage: A full-factorial experiment design can find all possible factor interactions where a simple experiment design may miss such interactions. Disadvantage: A full-factorial design has more experiments to run that a simple design.

Problem #7

Each sub-problem worth 3 points.

Answer the following questions regarding Markov models.

a) Precisely define the Markov property?

Past history is completely summarized by the specification of a current state.

b) What is a *P* matrix? What is a *Q* matrix? Briefly describe the key properties of each.

A *P* matrix is the matrix showing all state transition probabilities for a DMTC. A *Q* matrix is the matrix showing all state transition rates for a CMTC. For a P matrix each row sums to 1. For a *Q* matrix each row sums to 0 with each diagonal element being the negative sum of all other (non-diagonal) elements in the row.

c) What is the distribution of time between arrivals in a Poisson process? Show it.

The time between arrivals in a Poisson process is exponentially distributed. This can be shown as follows.

$$\Pr[T \le t] = 1 - \Pr[T > 0] = 1 - \Pr[\text{no arrival occurs in time } t]$$

$$= 1 - \frac{(\lambda t)^0}{0!} e^{-\lambda t} \text{ for 0 arrivals in a Poisson process in time } t$$

$$= 1 - e^{-\lambda t} \text{ which is the exponential distribution, so the time between arrivals follows an exponential distribution}$$

d) Convert the below *Q* matrix to a *P* matrix and then solve it (for steady state probabilities).

$$Q = \begin{bmatrix} -8 & 8 \\ 2 & -2 \end{bmatrix}$$

We can use uniformization, $P = I + (1/\gamma) \cdot Q$ where γ is the maximum absolute value in *Q* (6 in the case of the above example). So, we get

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \left(\frac{1}{8}\right) \cdot \begin{bmatrix} -8 & 8 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1/4 & 3/4 \end{bmatrix}$$

From $\pi = \pi \cdot P$ and $\sum \pi = 1$ we can write $\pi_0 = (1/4) \cdot \pi_1$ and $1 = \pi_0 + \pi_1$, which solves to $\pi_0 = 1/5$ and $\pi_1 = 4/5$.

Problem #8

4 pts for CTMC, 4 pts for Q matrix, 2 pts for set-up, 2 pts for correct numerical solution.

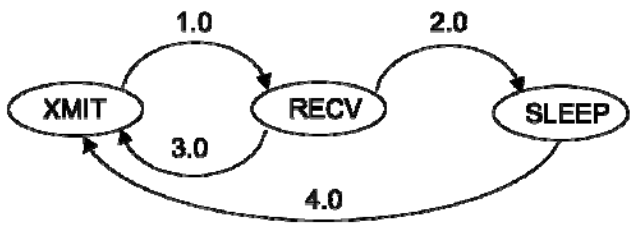
You are given a wireless device with three states, XMIT, RECV, and SLEEP. Let the power draw of each state be as follows:

- XMIT = 1 W
- RECV = 0.8 W
- SLEEP = 0.1 W

The system transitions between states at known rates (given below). What is the average power draw of this system?

- XMIT to RECV at rate = 1 transition per minute
- RECV to SLEEP at rate = 2 transitions per minute
- RECV to XMIT at rate = 3 transitions per minute
- SLEEP to XMIT at rate = 4 transitions per minute

As our first step, we draw the CTMC.



Let XMIT = state 0, RECV = state 1, and SLEEP = state 2

We then write the Q matrix:

$$Q = \begin{bmatrix} -1.0 & 1.0 & 0.0 \\ 3.0 & -5.0 & 2.0 \\ 4.0 & 0.0 & -4.0 \end{bmatrix}$$

We then write the three equations in three unknowns and solve for the steady state probabilities as:

$$\begin{aligned} -1.0\pi_0 + 3.0\pi_1 + 4.0\pi_2 &= 0 \\ 1.0\pi_0 - 5.0\pi_1 \qquad\qquad &= 0 \\ \pi_0 + \pi_1 + \pi_2 &= 1 \end{aligned}$$

Note that one of the equations is $\pi_0 + \pi_1 + \pi_2 = 1$ to break the linear dependence of the equations. We solve and get $\pi_0 = 10/13$, $\pi_1 = 2/13$, and $\pi_2 = 1/13$.

To determine the average power draw we multiply the steady state probabilities for each state by its power draw. That is,

$$\text{Average power draw} = 1.0\pi_0 + 0.8\pi_1 + 0.1\pi_2$$

which solves to 0.9 W.

Extra Credit +5 for difference. +5 for issue.

In "On the Modeling and Analysis of Computer Networks" (Kleinrock, *Proceedings of the IEEE*, August 1993) Kleinrock states that "Gigabits are indeed different" (compared to previous networks with kb/s and Mb/s data rates). Explain why Kleinrock states this. Describe at least one key issue with Gigabit networks as noted by Kleinrock.

Kleinrock says that the reason for the difference "has to do with the effect of latency due to the speed of light." The effect is that it is now latency, and not bandwidth, that dominates the time to deliver a file across long (WAN-scale) links. That is, the post-gigabit world is now latency, not bandwidth, limited.

A key issue is that of flow and congestion control. Feedback based schemes may be too sluggish where more bits may be "in flight" that the buffer at the receiver can store before a signal from the receiver to the sender to slow down is received. This suggest that proactive, versus reactive, flow control schemes need to be explored. One form of proactive flow control is rate-based flow control.

Humor

From: http://www.phdcomics.com

Hmm… maybe the non-thesis route is better after all. ☺

Appendix A – RDT FSMs from Kurose and Ross

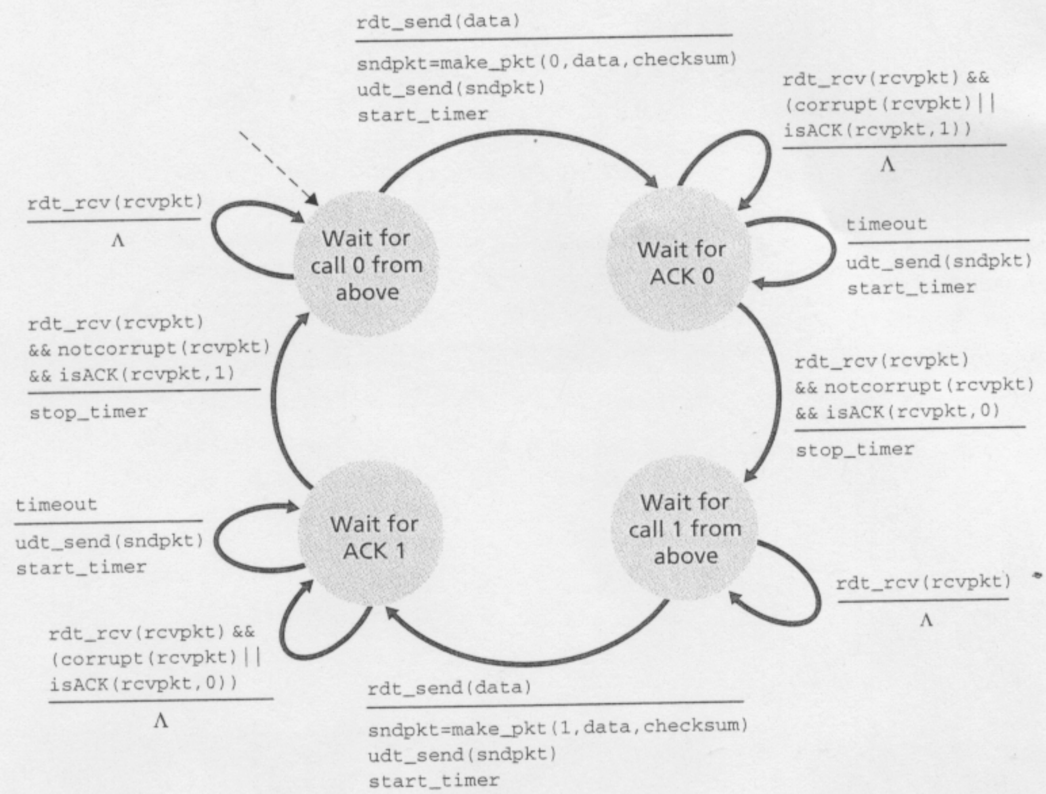


Figure 3.15 ♦ rdt3.0 sender

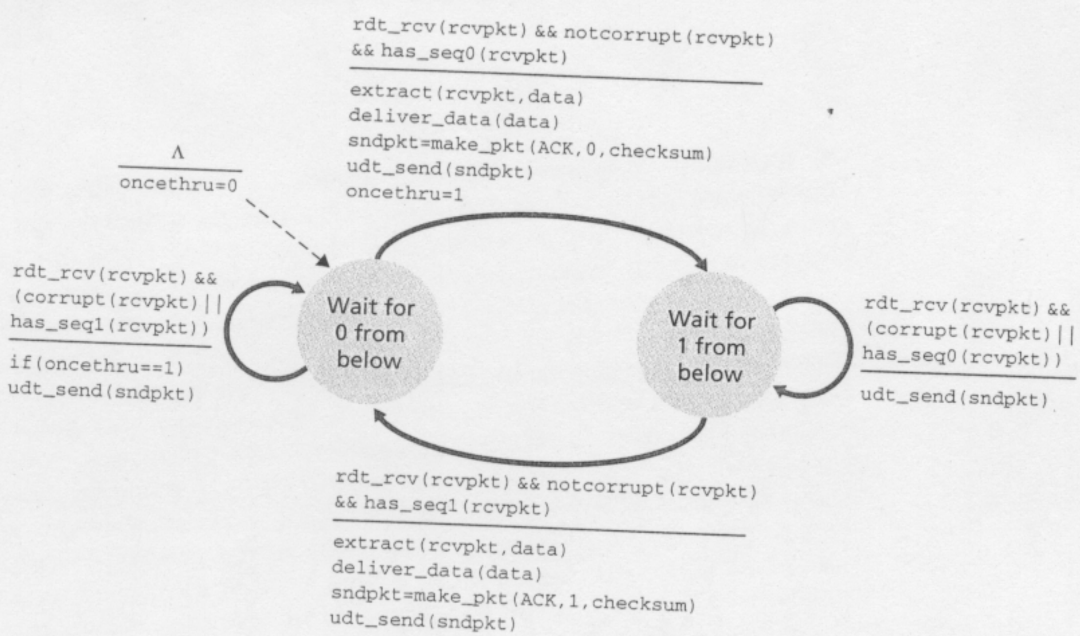


Figure 3.14 ♦ rdt2.2 receiver