PROBLEM 1

(a) Explain the meaning of *end-to-end paradigm* to achieving reliable data transport. How is reliability achieved (assuming the ARQ context)? Is reliable transport absolutely guaranteed? What are the main strengths of this approach? What are its main weaknesses?

(b) Assume you live in a world where transmission media are noise-free. That is, the physical layer assures that bits are transmitted without corruption between two hosts or routers connected to the same physical medium. Assume, furthermore, that devices such as hosts and routers never crash and all network software is guaranteed to be bug-free (a utopian world indeed). Since the physical layer and all other components are fail-proof, does this imply that ARQ is not needed at higher layers to achieve reliable transport in a WAN environment? Under what conditions is this assertion true and when is it false?

(c) As a continuation of part (b), at what other layers in the protocol stack can ARQ-based reliability be installed? Does this imply reliable end-to-end transport?

PROBLEM 2

(a) Assume the following bandwidth/data rate combinations for the transmission media *twisted pair* (copper wire), *coaxial cable* (also copper), and *optical fiber* of certain grade or quality, respectively: 4 Mbps/3 MHz, 500 Mbps/350 MHz, 2 Gbps/2 GHz. Assuming these depict numbers close to the fundamental limitation imposed by Shannon's 2nd Theorem (Channel Coding Theorem), estimate the signal-to-noise ratio underlying the three combinations. Is increasing the signal power (say, using more sophisticated signalling equipment) an effective way to increasing the throughput (or data rate) of a medium?

(b) Assume you have a *source*, $\langle \Sigma, \mathbf{p} \rangle$, given by the symbols $a \in \Sigma$ that it generates and the probability distribution $\mathbf{p}$ over the symbol set (or alphabet) Σ. By Shannon's First Theorem (Source Coding Theorem), we know that the average code length L_F of any code F is bounded below by the entropy of the source $H = -\sum_{a \in \Sigma} p_a \log p_a$. Given that there are n symbols (i.e., $n = |\Sigma|$), for what probability density $\mathbf{p}$ is entropy maximized? Compute the value. For what density is H minimal and what is its value?

(c) As a continuation of (b), what is the relation between Shannon's First Theorem to data compression? Does it say for what sources compression is most feasible? (This is more subtle than it looks.) How does this relate to the Kolmogorov complexity way of looking at randomness and compressibility? Give a unified explanation that puts everything into its proper place.

PROBLEM 3

Assume you have a noisy channel (e.g., wireless) where the probability of a bit flipping is given by $\varepsilon \geq 0$. Assume that the source is a binary source (i.e., $\Sigma = \{0, 1\}$) and $p_0 = 0.5$, $p_1 = 0.5$. Assume that to achieve reliable transmission you employ a straightforward redundancy scheme or code F where $F(a) = a^n$ for $a \in \{0, 1\}$ where "a^n" stands for the n-fold repetition (or string) of the symbol a. Here n is odd. Upon receiving a code word w of length n, decoding (i.e., F^{-1}) is performed by taking a majority vote of the 0s and 1s in w. That is, $F^{-1}(w) = 0$ iff the number of 0s is larger than the number of 1s in w (and vice versa for 1).

Assume the *stringency of reliable transmission*—the probability of a symbol being correctly decoded at the receiver—is given by a parameter $0 \leq \delta \leq 1$. When is $\delta = 1$ (perfect reliability) achievable? Is any redundancy ($n > 0$) needed? In general, to achieve reliability δ, how large does n need to be? Show your derivation. Compute n for $\delta = 0.999, 0.999999, 0.999999999$. What is the per-bit reliability (increase in reliability gained by incrementing n by 1) of this coding scheme? Would you say this code is efficient or desirable?

Assume that the channel in question is able to carry bits at 100 Mbps rate. For $\varepsilon = 0.000001$, to have on average one decoding error per 100 years (assuming the channel is utilized at full capacity throughout), what does n need to be?