momentum is less than P and the system is known to be inside a region of size R, then the phase space of the system must be divided into no more than $PR/\Delta P\Delta R = 2\pi PR/h$ distinguishable subintervals. This means that the number of distinguishable states n is bounded above by $2\pi PR/h$. Since for any particle, $P \leq E/c$, where E is the total energy of the system including the system's rest mass, with equality holding only if the system is moving at the speed of light, we have

$$I = \log_2 n \leq \frac{n}{\ln 2} \leq 2\pi \left(\frac{E}{c}\right)\left(\frac{R}{h \ln 2}\right) \leq \frac{2\pi ER}{\hbar c \ln 2}$$

which is the Bekenstein Bound (C.1). (Additional complications like particle substructure, and the fact that the system is in three dimensions rather than one are implicitly taken into account by the fact that $\log_2 N$ is very much less than N, for large N. As I said, the above derivation is nonrigorous.)

An upper bound to the information processing rate can be obtained [1] directly from the Bekenstein Bound by noting that the time for a state transition cannot be less than the time it takes for light to cross the sphere of radius R, which is $2R/c$. Thus

$$\dot{I} \leq \frac{I}{2R/c} \leq \frac{\pi E}{\hbar \ln 2} = 3.86262 \times 10^{51} \left(\frac{M}{1\text{kilogram}}\right) \text{bits/sec} \qquad (C.5)$$

where the dot denotes the proper time derivative. By inserting 100 kilograms for the value of M in inequality (C.5), we obtain an upper bound for the rate of change of state of a human being, $\dot{I}_{human}$:

$$\dot{I}_{human} \leq 3.86262 \times 10^{53} \text{states/sec} \qquad (C.6)$$

The significant digits in the RHS of inequalities (C.2), (C.3), (C.5), and (C.6) have to be taken with a grain of salt. The digits correctly express our knowledge of the constants c and $\hbar$. But the Bekenstein Bound is probably not the least upper bound to either I or $\dot{I}$; Schiffer and Bekenstein have recently shown [2] that the Bekenstein Bound probably overestimates both I and $\dot{I}$ by a factor of at least 100.

Strictly speaking, (C.5) only applies to a single communication channel [3], but it probably [4] applies even to multichannel systems if the need to merge the information from various channels is taken into account. However, if the latter is not taken into account, the number of channels is certainly limited by