1. We show that $n = R(3k, \ldots, 3k)^3$ (r copies) has the stated property (so the statement is true). Let $f : [n] \times [n] \times [n] \rightarrow [r]$ and for $1 \leq i < j < k \leq n$ set $g(i, j, k) = f(i, j, k)$. Thus $g : (\binom{n}{3}) \rightarrow [r]$ and by our choice of $n$ there is some $D = \{i_1, \ldots, i_{3k}\} \subseteq [n]$ with $g$ constant on $\binom{D}{3}$. Now take $A = \{i_1, \ldots, i_k\}$, $B = \{i_{k+1}, \ldots, i_{2k}\}$ and $C = \{i_{2k+1}, \ldots, i_{3k}\}$.

2.(a) Given $\delta, l$, take $L_0$ as in Szemerédi’s Theorem (so for any $L > L_0$, any $A \subseteq [L]$ of size at least $\delta L$ contains an $l$-term A.P.). For $x \in X$ set

$$A_x = \{i \in [L] : x \in Y_i\} \ (x \in X).$$

It’s enough to show that there’s some $x \in X$ for which $|A_x| \geq \delta|L|$: our choice of $L$ then guarantees that $A_x$ contains an A.P. $a, a+d, \ldots, a+(l-1)d$, which gives the desired $Y_i$’s. But $\delta|L|$ is a lower bound on the average of the $|A_x|$’s:

$$|X|^{-1} \sum_{x \in X} |A_x| = |X|^{-1} \sum_{i=1}^{L} |Y_i| \geq |X|^{-1} |L| \delta |X| = \delta |L|.$$

(b) We show $N_0 = L_0(\delta/2, k)$ has the desired property. Let $N > N_0$, $A \subseteq [N]$ and $|A| > \delta N$. Set $X = [2N]$ and $Y_i = A + i$ for $i \in [N]$. Then $Y_i \subseteq X$ and $|Y_i| = |A| > (\delta/2)|X|$; so, since $N > N_0$, there are $a, d \in \mathbb{P}$ so that $Y_a \cap Y_{a+d} \cap \cdots \cap Y_{a+(l-1)d} \neq \emptyset$. But for any $x$ in this intersection, $\{x - a, x - (a + d), \ldots, x - (a + (k - 1)d)\}$ is a $k$-term A.P. contained in $A$.

3. Let $X = \{A \subseteq [n] : n/2 - K \sqrt{n} < |A| < n/2 + K \sqrt{n}\}$, with the constant $K$ chosen so $|2^{[b]} \setminus X| < \varepsilon 2^{n-1}$. Then $|S \cap X| > \varepsilon 2^{n-1}$. Let $C$ be the collection of maximal chains in $X$. Then $|C| < 2K \sqrt{n}$ for each $C \in C$, while

$$\frac{1}{|C|} \sum_{C \in C} |C \cap S| = \sum_{x \in S \cap X} \frac{|C : x \in C \in C|}{|C|} \geq \frac{|S \cap X|}{\frac{n}{|n/2|}} > \frac{\varepsilon}{2} \sqrt{n}.$$

Thus there is a $C \in C$ with

$$|C \cap S| > \frac{\varepsilon}{2} \sqrt{n} > \frac{\varepsilon}{4K} |C|,$$

and we may apply Szemerédi’s Theorem to obtain (for large enough $n$) an arithmetic progression in $C$. 

1
Alternate (using something from last semester): Let $C$ be a symmetric chain decomposition of $2^{[n]}$. Since $|C| = O(2^n/\sqrt{n})$, there are at most $o(2^n)$ elements of $S$ contained in chains of length at most (say) $n^{1/4}$; so there is some $C \in C$ with $|C| > n^{1/4}$ and $|S \cap C| > \varepsilon |C|/2$, etc.

4. Let $A_0 = \{ x \in A : d_B(x) > 3n \}$. We know $|A_0| \geq (1 - \varepsilon) n > (2\beta - 1)n$ (the second inequality holding because $\beta + \varepsilon \leq 1$). So $A_0$ contains an edge $xy$. But $|N_B(x) \cap N_B(y)| > (2\beta - 1)n$, so there is an edge $zw$ in $N_B(x) \cap N_B(y)$, and $G\{x, y, z, w\} = K_4$.

5. Say $x_k \in X_i$ is good if $\min\{|\{l < k : x_l \in X_i\}|, |\{l > k : x_l \in X_i\}|\} > \varepsilon m$, define “$y_k$ good” similarly, and say $k$ is good if both $x_k$ and $y_k$ are.

Observation: If $k$ is good, $x_k \in X_i$ and $y_k \in Y_j$, then $(X_i, Y_j)$ is $(\varepsilon)$-irregular.

The number of good $k$’s is at least $(1 - 4\varepsilon)n$ (why?), and $X_i$ is in an irregular pair if there’s at least one good $k$ with $x_k \in X_i$; so since any $X_i$ contains at most $(1 - 2\varepsilon)m$ such $x_k$’s (the trivial $m$ would also be okay here), the number of $X_i$’s in irregular pairs is at least $\frac{1 - 4\varepsilon}{1 - 2\varepsilon} m = \frac{1 - 4\varepsilon}{1 - 2\varepsilon} t$.

For the bonus (the answer is yes): Start with partitions $X = U_1 \cup \cdots \cup U_t$, $Y = V_1 \cup \cdots \cup V_t$, where the $U_i$’s and $V_i$’s are intervals written in the natural order, but with $|U_i| = |V_i| = m$ and all other blocks of size $(t - 1)m/t$. Now revise to $X_1, \ldots, X_t$ and $Y_2, \ldots, Y_{t+1}$, where $X_i = U_i, Y_{t+1} = V_{i+1}$: for $i \geq 2$, $X_i$ is $U_i$ plus $m/t$ elements of $U_{i+1}$; and for $i \leq t$, $Y_i$ is $V_i$ plus $m/t$ elements of $V_1$. (And show this works for $t > 2\varepsilon^{-2}$, with the only irregular pairs being $(X_i, Y_i)$, $2 \leq i \leq t$.)

6. First a convenient (though avoidable) observation (why is it true?):

if $(X, Y)$ is $\varepsilon$-irregular with density $d$, then there are $X' \subseteq X$ and $Y' \subseteq Y$ with $d(X', Y') \notin (d - \varepsilon, d + \varepsilon)$ and $|X'| = \varepsilon |X|, |Y'| = \varepsilon |Y|$

(where, as usual, we pretend all large numbers are integers).

Lemma. If $(X, Y)$ is $\varepsilon$-irregular with density $d$, then there are $X_0 \subseteq X$ and $Y_0 \subseteq Y$ with $|X_0| \geq \varepsilon |X|$, $|Y_0| \geq \varepsilon |Y|$ and $d(X_0, Y_0) > d + \varepsilon^3$. (Then start with any equipartition $V = A \cup B$ and apply the lemma at most $\varepsilon^{-3}$ times to obtain an $\varepsilon$-regular pair with part sizes at least $\varepsilon^{-3} n/2$.)

Proof of lemma. Let $X', Y'$ be as above. If $d(X', Y') > d + \varepsilon$, take $(X_0, Y_0) = (X', Y')$. If $d(X', Y') < d - \varepsilon$, then (with $X' = X \setminus X', Y' = Y \setminus Y'$)

\[
d = \varepsilon^2 d(X', Y') + \varepsilon (1 - \varepsilon)[d(X', Y) + d(X, Y')] + (1 - \varepsilon)^2 d(X', Y') \\
< \varepsilon^2 (d - \varepsilon) + \varepsilon (1 - \varepsilon)[d(X', Y) + d(X, Y')] + (1 - \varepsilon)^2 d(X', Y'),
\]

(1)
implying that (at least) one of the densities on the r.h.s. of (1) is greater than 
\[ [(1 - \varepsilon^2)d + \varepsilon^3]/(1 - \varepsilon^2) > d + \varepsilon^3. \]

7. We work with the \( \delta \) version and may assume \( \delta \) is fairly small. Fix \( \varepsilon \ll \delta^2 \) and let \( t_0 = \varepsilon^{-1} \) and \( T = T(\varepsilon, t_0) \) (as in the Regularity Lemma). We may specify a triangle-free \( G \) on \( n \) by specifying:

1. a partition \( V_1 \cup \cdots \cup V_{t_0} \) (with \( t \leq T \)) that is \((\varepsilon, \delta)\)-regular for \( G \);
2. the pairs \( \{V_i, V_j\} \) that are either irregular or have density less than \( \delta^2 \);
3. the edges of \( G \) that are either contained in some \( V_i \) or belong to one of the pairs specified in (2);
4. the remaining edges of \( G \) (those belonging to “dense” regular pairs).

(Of course there may be many such specifications for a given \( G \), but this is fine: we’re just bounding the number of \( G \)’s by the number of specifications.)

The numbers of possibilities in (1) and (2) are bounded by \( T \cdot T^n = \exp[O(n)] \) and \( O(1) \) respectively (of course the constants are huge), and we may, for example, bound the number of choices in (3) by the total number of sets of up to \( (\varepsilon + \delta^2)n^2 \) pairs from \( [n] \), which is less than \( \exp[2\delta^2 \log(\varepsilon/\delta^2)n^2] \) (say), which is fine since \( \delta^2 \log(\varepsilon/\delta^2) \) is much smaller than \( \delta \).

Now the point: since \( G \) is triangle-free, the Counting Lemma (and our choice of \( \varepsilon \)) implies that there cannot be (distinct) \( i, j, k \) with each of the pairs from \( \{V_i, V_j, V_k\} \) dense and regular. So by Mantel’s Theorem the number of pairs \( \{V_i, V_j\} \) in (4) is less than \( t^2/4 \), whence the number of possibilities for this step is less than \( \exp_2[(k^2/4)(n/k)^2] = \exp_2[n^2/4] \).

8. (Due to Ruzsa and Szemerédi and the original reason for the Triangle Removal Lemma.) Let \( G \) be the graph on \( V(\mathcal{H}) \) with \( xy \in E(G) \) iff some edge of \( \mathcal{H} \) contains \( \{x, y\} \). By \((\ast)\): (i) the only triangles in \( G \) are those gotten (in the obvious way) from edges of \( \mathcal{H} \), and (now using (i)) (ii) no edge of \( G \) belongs to more than two triangles. Thus \( |\mathcal{H}| = \tau(G) \leq 2|G|/3 = o(n^3) \), so by the TRL, (recall \( \rho \) is distance from triangle-free) \( |\mathcal{H}| \leq 2\rho(G) < o(n^2) \).

9. (A giveaway provided for use in 10(a).) With usage as in class and sums over \( \{x, y, z\} \), we may write \( H(X, Y, Z) = H(Y) - H(X, Y) - H(X, Y, Z) \) as

\[
\sum p(x, y, z) \log \frac{p(x, y)p(y, z)}{p(x, y, z)p(y)} \leq \log \left[ \sum p(x, y, z) \frac{p(x, y)p(y, z)}{p(x, y, z)p(y)} \right] \leq \log \left[ \sum \frac{p(x, y)p(y, z)}{p(y)} \right] = \log \left[ \sum p(y)p(x|y)p(z|y) \right] = \log 1 = 0.
\]
10. (a) True, since \( f(i, i + 1, i + 2, i + 3) + f(i + 1, i + 2) \leq f(i, i + 1, i + 2) + f(i + 1, i + 2, i + 3) \) (instance of \( H(X, Y, Z) + H(Y) \leq H(X, Y) + H(Y, Z) \)).

(b) False; here’s a nice symmetric example, though there are others that are even easier: let \( Y_1, \ldots, Y_5 \) be independent fair coin flips and \( X_i = (Y_i, Y_{i+1}) \).