Name _____ Student No. _____

No aids allowed. Answer all questions on test paper. Use backs of sheets for scratch work. ***Show all your work; there will be no credit for answers without a justification.***

Total Marks 100: four questions, each worth 25.

1. A sequence of positive integers $r_1, r_2, \ldots, r_n$ is *super-increasing* if $r_{i+1} \geq 2r_i$ for all $1 \leq i \leq n-1$.

   Show that if a sequence is super-increasing, then $r_{i+1} > \sum_{j=1}^{i} r_i$ for all $1 \leq i \leq n-1$.

   SOLUTION: We show this by induction on i. When $i = 1$ then $r_2 \geq 2 \cdot r_1 > r_1$, where the last strict inequality follows from the fact that r_1 is a positive integer.

   For the induction step, assume that we have $r_{i+1} > \sum_{j=1}^{i} r_i$, and consider

   $$r_{i+2} \geq 2r_{i+1} = r_{i+1} + r_{i+1} > r_{i+1} + \sum_{j=1}^{i} r_i = \sum_{j=1}^{i+1} r_i$$

   where the last strict inequality follows by the induction hypothesis.

2. Let $M_1, M_2, \ldots, M_n, C$ be a subset sum problem where the weights $M_1, M_2, \ldots, M_n$ are super-increasing. Assume that a solution $x = x_1 x_2 \ldots x_n \in \{0,1\}^n$ exists (i.e., $C = \sum_{i=1}^{n} x_i M_i$); describe an efficient algorithm to compute it, and then show that the algorithm is correct.

SOLUTION: The algorithm is the following

1. $S \longleftarrow C$
2. for $i = n, \ldots, 1$
3. if $S \geq M_i$ then
4. $x_i \longleftarrow 1$ and $S \longleftarrow S - M_i$
5. else
6. $x_i \longleftarrow 0$

We are guaranteed that a solution exists; we call it $y = y_1 y_2 \ldots y_n \in \{0,1\}^n$. To show correctness we are going to show by induction on $i = 0, 1, \ldots, n-1$ that $y_{n-i} = x_{n-i}$ (here x is the output of the algorithm).

The basis case is $i = 0$, so we want to show that $y_n = x_n$. We consider two cases: case 1, $y_n = 1$ so M_n is in the y-solution, so in particular $M_n \leq C$, so the first time the algorithm executes line 3 we have that $S \geq M_n$, and so x_n is also set to 1. Case 2, $y_n = 0$ so M_n is not in the y-solution. So we know that $C = \sum_{i=1}^{n} y_i M_i < M_n$, and so line 3 of the algorithm will not be true, and so $x_n = 0$.

Now the induction step. Suppose that the claim is true for $i = 0, 1, \ldots, k-1$, i.e., $y_n = x_n, y_{n-1} = x_{n-1}, \ldots, y_{n-k+1} = x_{n-k+1}$. Now we show that $y_{n-k} = x_{n-k}$ as well. To this end we essentially repeat the basis case. Case 1, $y_{n-k} = 1$ so M_{n-k} is in the y-solution. Since at this point $S = \sum_{i=0}^{k-1} x_{n-i} M_{n-i} = \sum_{i=0}^{k-1} y_{n-i} M_{n-i}$ it follows that $M_{n-k} \leq S$, and so line 3 is true and $x_{n-k} = 1$. Case 2, $y_{n-k} = 0$ so M_{n-k} is not in the y-solution. We know that $\sum_{i=1}^{n-k-1} y_i M_i < M_{n-k}$, which means that $M_{n-k} > C - \sum_{i=0}^{k-1} y_{n-i} M_{n-i} = S$, so line 3 is false, and hence $x_{n-k} = 0$.

3. Explain how to construct a cryptosystem based on a super-increasing subset sum problem.

SOLUTION: Alice starts with a super-increasing sequence $r_1, r_2, \ldots, r_n$, and chooses two large secret keys A, B satisfying $B > 2r_n$ and $\gcd(A, B) = 1$. She now "scrambles" the super-increasing sequence to get $M_i = Ar_i \pmod{B}$, where $M_i \in \mathbb{Z}_B$. The sequence thus obtained, $M_1, M_2, \ldots, M_n$ is Alice's public key.

To send a plaintext message $x \in \{0,1\}^n$, Bob computes and sends to Alice the cyphertext $S = \sum_{i=1}^{n} x_i M_i$. To decrypt S, Alice first computes $S' = A^{-1}S \pmod{B}$, where $S' \in \mathbb{Z}_B$. Then Alice solves the subset sum problem for S' using the algorithm in the previous question.

The reason why the decryption works is the following:

$$S' = A^{-1}S = A^{-1}\sum_{i=1}^{n} x_i M_i = A^{-1}\sum +i = 1^n x_i A r_i = \sum_{i=1}^{n} x_i r_i \pmod{B}.$$

4. Let $v_1, v_2, \ldots, v_n$ be a basis for a vector subspace of $\mathbb{R}^m$. Consider the Gram-Schmidt algorithm:

1. $v_1^* \longleftarrow v_1$
2. for $i = 2, \ldots, n$
3. for $j = 1, \ldots, i-1$
4. $\mu_{ij} \longleftarrow (v_i \cdot v_j^*)/\|v_j^*\|^2$
5. $v_i^* \longleftarrow v_i - \sum_{j=1}^{i-1} \mu_{ij} v_j^*$

Show that the output $\{v_1^*, v_2^*, \ldots, v_n^*\}$ is an orthogonal basis with the property that $\text{span}\{v_1, v_2, \ldots, v_n\} = \text{span}\{v_1^*, v_2^*, \ldots, v_n^*\}$.

SOLUTION: To show that the set $\{v_1^*, v_2^*, \ldots, v_n^*\}$ is orthogonal we have to show that the dot-product of any two distinct vectors (in it) is zero. We do this by induction on $i = 1, 2, \ldots, n$. The basis case is trivial since when $i = 1$ the set consists of $\{v_1^*\}$ and there are no distinct pairs. Suppose now that $\{v_1^*, v_2^*, \ldots, v_i^*\}$ is orthogonal; it is sufficient to show that the dot-product of v_{i+1}^* with any vector in that set is zero. Indeed, for $k = 1, \ldots, i$,

$$v_{i+1}^* \cdot v_k^* = \left(v_{i+1} - \sum_{j=1}^{i} \mu_{(i+1)j} v_j^* \right) \cdot v_k^* = v_{i+1} \cdot v_k^* - \mu_{(i+1)k} \|v_k^*\|^2 = 0$$

To show the statement about the spans, note that from line 5 it is clear that v_i is in the span of $\{v_1^*, v_2^*, \ldots, v_i^*\}$. We prove the other inclusion by induction on i as follows: $\text{span}\{v_1^*, \ldots, v_i^*\} \subseteq \text{span}\{v_1, \ldots, v_i\}$. The basis case, $i = 1$, follows trivially from line 1 of the algorithm. The induction step follows immediatelly from line 5.