Proof of Theorem 4.7.4

The following proof is for the case of a *nondecreasing* monotonic optimal policy (the proof is very similar for a nonincreasing optimal policy):

As discussed in class, from Lemma 4.7.1 it is sufficient to prove that $v_t(s,a)$ is superadditive. We can prove that $v_t(s,a)$ is superadditive whenever the stated conditions (1)-(5) hold as follows. By condition (2) and the definition of superadditivity, for any two given states s^-, s^+ such that $s^- \leq s^+$ and for all k in the set of states S:

$$\sum_{j=0}^{\infty} p_t(j\,|s^-,a^-) + p_t(j\,|s^+,a^+) \geq \sum_{j=0}^{\infty} p_t(j\,|s^-,a^+) + p_t(j\,|s^+,a^-)$$

Using conditions (1), (3), and (5), by proposition 4.7.3 $v_t(s)$ is nondecreasing in s for all t. Applying Proposition 4.7.2 yields:

$$\sum_{j=0}^{\infty} (p_t(j\,|s^-,a^-) + p_t(j\,|s^+,a^+))v_t(j) \geq \sum_{j=0}^{\infty} (p_t(j\,|s^-,a^+) + p_t(j\,|s^+,a^-)v_t(j)$$

Thus it follows that $\sum_{j=0}^{\infty} p_t(j|s,a)v_t(j)$ is superadditive. From condition (3) and the fact that the sum of superadditive functions is superadditive it follows that $v_t(s,a)$ is superadditive. Thus by Lemma 4.7.1 the optimal policy is nondecreasing and the proof is complete.