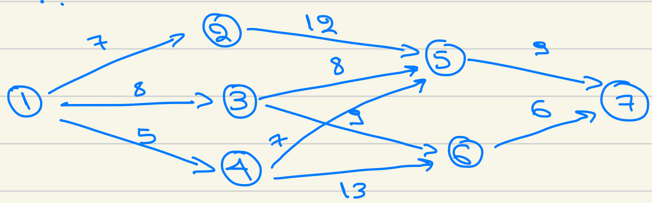
Ενότητα 6: Προβλήματα Δυναμικού προγραμματισμού

Ασχολούμαστε με προβλήματα που οι αποφάσεις λαμβάνονται σε στάδια. Σε κάθε στάδιο πρέπει να πάρω μία απόφαση.

Κεντρική ιδέα: Σπάμε το πρόβλημα σε υποπροβλήματα (στάδια) και σε κάθε στάδιο έχουμε μία μεταβλητή απόφασης.

Παράδειγμα 1: (Πρόβλημα ελάχιστης διαδρομής)

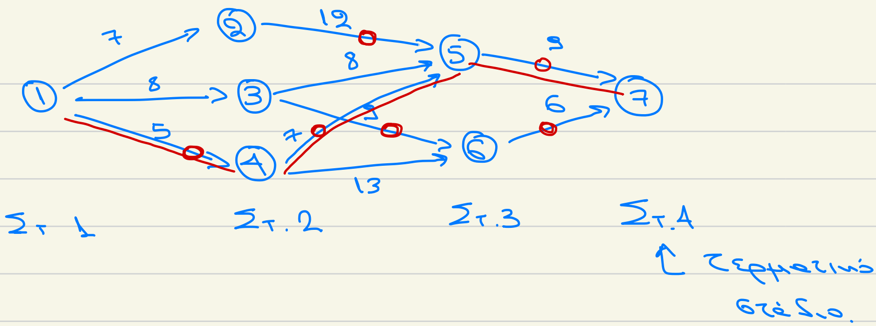
Ψάχνουμε την ελάχιστη διαδρομή από την πόλη 1 στην πόλη 7.



Λύση:

Παρατηρούμε ότι:

- Κάθε διαδρομή αποτελείται από 3 βήματα, άρα θα πάρουμε 3 διαδοχικές αποφάσεις.
- Σε κάθε βήμα, ανάλογα με την πόλη που βρίσκομαι (δηλαδή ανάλογα με την κατάσταση στην οποία βρίσκομαι) έχω συγκεκριμένες δυνατές αποφάσεις.
- Κάθε απόφαση έχει κάποιο κόστος (άμεσο) και οδηγεί σε συγκεκριμένη επόμενη κατάσταση.



$v(x)$ = μήκος ελάχιστης διαδρομής από την πόλη x στην πόλη 7.

Θέλω να βρω το $v(1)$.

<u>Στάδιο 3</u>

βέλτιστη απόφαση όταν είμαι στην πόλη 5, στο στάδιο 3

$v(5) = 9 \quad , \quad a^*(5) = 7$

$v(6) = 6 \quad , \quad a^*(6) = 7$

<u>Στάδιο 2</u>

$v(2) = 12 + v(5) = 12 + 9 = 21 \quad , \quad a^*(2) = 5$

$$v(3) = \min\{\underbrace{8 + v(5)}_{\text{απόφ } 5}, \underbrace{9 + v(6)}_{\text{απόφ } 6}\} =$$

$$= \min\{8 + 9, 9 + 6\} = \min\{17, 15\} = 15,$$

$a^*(3) = 6$

$$v(4) = \min\{\underbrace{7 + v(5)}_{\text{απόφ } 5}, \underbrace{13 + v(6)}_{\text{απόφ } 6}\} =$$

$$= \min\{7 + 9, 13 + 6\} = \min\{16, 19\} = 16,$$

$a^*(4) = 5$

Στάδιο 1

$$v(1) = \min\{\underbrace{7+v(2)}_{\text{απόφ } 2},\ \underbrace{8+v(3)}_{\text{απόφ } 3},\ \underbrace{5+v(4)}_{\text{απόφ } 4}\} =$$

$$= \min\{7+21,\ 8+15,\ 5+16\} =$$

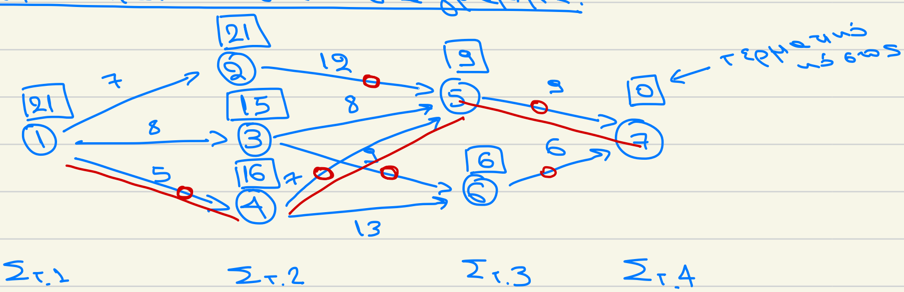
$$= \min\{28,\ 23,\ 21\} = 21 \qquad a^*(1) = 4$$

Μήκος ελάχιστης διαδρομής από την πόλη 1 μέχρι το τέλος είναι 21 και η ακολουθία βέλτιστων αποφάσεων είναι

$$1 \longrightarrow 4 \longrightarrow 5 \longrightarrow 7$$

Στ.1 Στ.2 Στ.3 Στ.4

Επίλυση πάνω σε διάγραμμα:

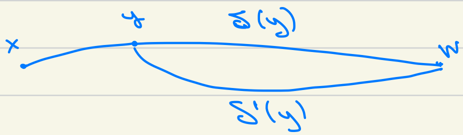


Η ελάχιστη διαδρομή έχει μήκος 21 και η βέλτιστη ακολουθία αποφάσεων είναι:

$$1 \longrightarrow 4 \longrightarrow 5 \longrightarrow 7$$

Στ.1 Στ.2 Στ.3 Στ.4

<u>Αρχή βελτιστοποίησης Bellman</u>



Αν $\delta(x) = ((x,y), \delta(y))$ είναι βέλτιστη από το x στο w, τότε η $\delta(y)$ θα είναι βέλτιστη από το y στο w.

<u>Απόδειξη</u>

Έστω ότι η $\delta(y)$ δεν είναι βέλτιστη από το y στο w και υπάρχει $\delta'(y)$ με μικρότερο κόστος. Τότε η $\delta'(x) = ((x,y), \delta'(y))$ θα έχει μικρότερο κόστος από την $\delta(x)$. Άτοπο.

Συμβολισμοί - Ορισμοί

Σε ένα πρόβλημα δυναμικού προγραμματισμού πεπερασμένου χρονικού ορίζοντα με ντετερμινιστική δυναμική ορίζουμε:

- **Στάδια:** $t = 1, 2, 3, \ldots, N, N+1$
  N = μήκος χρονικού ορίζοντα (# διαδοχικών αποφάσεων)
  $N+1$ = τερματικό στάδιο (δεν παίρνω απόφαση)
- **Καταστάσεις:** x_t : κατάσταση στο στάδιο t
  (στη αρχή του σταδίου t πριν πάρω απόφαση για το στάδιο t)
  S_t = χώρος καταστάσεων στο στάδιο t.
  (Θα ασχοληθούμε μόνο με προβλήματα με πεπερασμένο σύνολο καταστάσεων).
- **Αποφάσεις:** a_t : απόφαση στο στάδιο t.
  $D_t(x_t)$ = σύνολο δυνατών αποφάσεων στο στάδιο t όταν βρίσκομαι στην x_t.
  (Θα ασχοληθούμε με π.δ.π. με πεπερασμένο σύνολο δυνατών αποφάσεων)
- **Δυναμική συστήματος:** Ο μηχανισμός που ορίζει την κατάσταση στο επόμενο στάδιο σαν αποτέλεσμα της τωρινής κατάστασης και της απόφασης.

$$x_{t+1} = g_t(x_t, a_t)$$

- **Άμεσο κόστος/κέρδος:** $c_t(x_t, a_t)$

- **Τερματικό κόστος/κέρδος:** $\hat{c}(x_{N+1})$

  κόστος ή κέρδος επειδή στο τέλος του χρονικού ορίζοντα βρίσκομαι στην κατάσταση x_{N+1}.

- **Εξισώσεις βελτιστοποίησης (Bellman)**

$$v(t, x_t) = \min/\max_{a_t \in D_t(x_t)} \left\{ c_t(x_t, a_t) + v(t+1, g_t(x_t, a_t)) \right\}, \quad t = 1, 2, \dots, N$$

με τερματικές συνθήκες

$$v(N+1, x_{N+1}) = \hat{c}(x_{N+1})$$

όπου

$v(t, x_t)$ = ελάχιστο κόστος ή μέγιστο κέρδος από το στάδιο t μέχρι το τέλος αν στην αρχή του σταδίου t το πρόβλημα βρίσκεται στην κατάσταση x_t.

Παράδειγμα 2:

Ένα φοιτητής μπορεί να επιλέξει (έχει) 7 μαθήματα από 4 κατηγορίες, με τουλάχιστον ένα μάθημα από κάθε κατηγορία. Ο παρακάτω πίνακας δίνει ένα σκορ που μετράει τη γνώση που έχει αποκτήσει παίρνοντας κάποιον αριθμό μαθημάτων από κάθε κατηγορία.

Κατηγορία (t)	# μαθημάτων (a_t): 1	2	3	4
1	25	50	60	80
2	20	70	90	100
3	40	60	80	100
4	10	20	30	40

$= R(t, a_t)$

Να βρεθεί πόσα μαθήματα πρέπει να επιλέξει από κάθε κατηγορία για να μεγιστοποιήσει το σκορ.

Λύση.

Μοντελοποίηση ως π.δ.π:

- **Στάδια:** $t = 1, 2, 3, 4, 5$ → Τερματικό στάδιο

  $N = 4$

  Στο στάδιο t αποφασίζω πόσα μαθήματα θα πάρω από την κατηγορία t.

- **Κατάσταση:** x_t = # μαθημάτων που απομένουν στην αρχή του σταδίου t

  $x_1 = 7$

- **Αποφάσεις:** a_t = # μαθημάτων που θα πάρω από την κατηγορία t (στο στάδιο t)

$$D_t(x_t) = \{1, \dots, x_t - (4-t)\}$$

$$D_2(5) = \{1, 2, \underbrace{3}_{5-2}\}$$

στην αρχή του σταδίου 2 απομένουν 5 μαθήματα.
Υποχρεωτικά στο στάδιο 3 και 4 πρέπει να πάρω από ένα μάθημα.

$x_3 = 3$ $\quad$ $x_4 = 3 - 2 = 1$

$a_3 = 2$

- <u>Δυναμική συστήματος</u>: $x_{t+1} = g_t(x_t, a_t)$

  $x_{t+1} = x_t - a_t$

- <u>Άμεσο κέρδος</u>: $c_t(x_t, a_t) = R(t, a_t)$

  ≠ μαθημάτων από κατηγορία t

  κατηγορία

- <u>Τερματικό κέρδος</u>: $\hat{c}(x_5) = 0$

  Αν έπρεπε να πάρει ακριβώς 7 μαθήματα:

$$\hat{c}(x_5) = \begin{cases} 0, & \text{αν } x_5 = 0 \\ -\infty, & \text{αν } x_5 \neq 0 \end{cases}$$

- <u>Εξισώσεις βελτιστοποίησης</u>

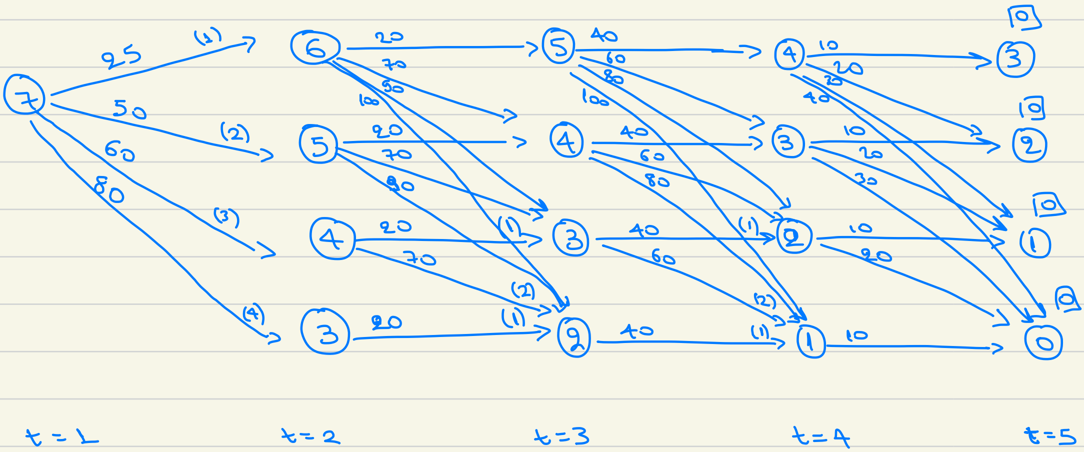
  $v(t, x_t)$ = μέγιστο κέρδος από το στάδιο t μέχρι το τέλος αν στην αρχή του σταδίου t η κατάσταση είναι η x_t.

$$v(t, x_t) = \max_{a_t \in D_t(x_t)} \{ c_t(x_t, a_t) + v(t+1, g_t(x_t, a_t)) \},$$

$t = 1, 2, 3, 4$

με τερματικές συνθήκες

$$v(5, x_5) = \hat{c}(x_5) = 0$$


t = 1
t = 2
t = 3
t = 4
t = 5