

עבודה 1- מכונות נבונות- למידה וקבלת החלטות

(1) הגדרת הבעיה כבעיית MDP:

- מצבים:

מצב מוגדר ע"י:

<שעה, מצב החולה אצל הרופא, זמן שנותר לחולה בבית חולים, האם החולה האחרון שאובחן שרד >

- מצב התחלתי < false, 0, Unknown, 9 >

- מצב ביניים < *, *, *, (10 - 14) > חוץ ממצב סופי.

- מצב סופי < *, *, Unknown, 14 >

ערכי שדות אפשריים:

מצב חולה אצל הרופא: Flu, Ebola, Cough, Unknown.

שעה: 9-14.

זמן שנותר לחולה בבית החולים: 0-2.

האם החולה האחרון שאובחן שרד: true/false.

- פעולות:

A0 – אבחון.

A1 – שליחה הביתה.

A2 – שליחה לבית החולים.

ממצבים < *, *, Unknown, 9-13 > הפעולה האפשרית היחידה היא אבחון-A0.

ממצבים < false, 0, not(Unknown), 10-14 > ניתן לבחור אחת מבין הפעולות:

A1 – שליחה הביתה.

A2 – שליחה לבית החולים.

ממצבים < false, 1, not(Unknown), 10-14 > הפעולה האפשרית היחידה היא שליחה הביתה-A1.

ממצבים סופיים < *, *, Unknown, 14 > אין פעולות אפשריות.

- Reward function:

- 1 + אם ערך השדה "האם החולה האחרון שאובחן שרד" true.

- 0 בכל מצב אחר.

- Transition function:

מעברים לפעולה A0:

$$T_{A0} < *, t, Unknown, z > = < false, \max\{t-1, 0\}, Flu, z+1 > \mid p=0.8$$

$$T_{A0} < *, t, Unknown, z > = < false, \max\{t-1, 0\}, Ebola, z+1 > \mid p=0.1$$

$$T_{A0} < *, t, Unknown, z > = < false, \max\{t-1, 0\}, Cough, z+1 > \mid p=0.1$$

מעברים לפעולה A1:

$$T_{A1} < \text{false}, t, \text{Flu}, z > = < \text{true}, t, \text{Unknown}, z >$$

$$T_{A1} < \text{false}, t, \text{Cough}, z > = < \text{true}, t, \text{Unknown}, z > \mid p=0.5$$

$$T_{A1} < \text{false}, t, \text{Cough}, z > = < \text{false}, t, \text{Unknown}, z > \mid p=0.5$$

$$T_{A1} < \text{false}, t, \text{Ebola}, z > = < \text{false}, t, \text{Unknown}, z >$$

מעברים לפעולה A2:

$$T_{A2} < \text{false}, 0, \text{Flu}, z > = < \text{true}, 1, \text{Unknown}, z > \mid p=0.5$$

$$T_{A2} < \text{false}, 0, \text{Flu}, z > = < \text{true}, 2, \text{Unknown}, z > \mid p=0.5$$

$$T_{A2} < \text{false}, 0, \text{Cough}, z > = < \text{true}, 1, \text{Unknown}, z > \mid p=0.5$$

$$T_{A2} < \text{false}, 0, \text{Cough}, z > = < \text{true}, 2, \text{Unknown}, z > \mid p=0.5$$

$$T_{A2} < \text{false}, 0, \text{Ebola}, z > = < \text{true}, 1, \text{Unknown}, z > \mid p=0.5 \cdot 0.25 = 0.125$$

$$T_{A2} < \text{false}, 0, \text{Ebola}, z > = < \text{true}, 2, \text{Unknown}, z > \mid p=0.5 \cdot 0.25 = 0.125$$

$$T_{A2} < \text{false}, 0, \text{Ebola}, z > = < \text{false}, 1, \text{Unknown}, z > \mid p=0.5 \cdot 0.75 = 0.375$$

$$T_{A2} < \text{false}, 0, \text{Ebola}, z > = < \text{false}, 2, \text{Unknown}, z > \mid p=0.5 \cdot 0.75 = 0.375$$

(2) תוצאות הרצת value iteration:

Optimal policy:

State	Optimal policy
< *, 0, Unknown, 9-13 >	A0 (diagnose)
< false, 0, Flu, 10-14 >	A1 (home)
< false, 1, Flu/Cough/Ebola, 11-14 >	A1 (home)
< false, 0, Cough/Ebola, 10-14 >	A2 (hospital)

המספר הממוצע של אנשים שישרדו בpolicy זה הוא 4.597.

(3) מידול בעיה A: הרחבת שעות פעילות המרפאה.

השינוי היחיד בעקבות הרחבת שעות הפעילות של המרפאה הוא שכעת השעה במצבי הסיום לא תהיה 14, אלא שעת הסיום החדשה.

לדוגמא, אם שעת הסיום החדשה היא 8pm, סוגי המצבים יהיו:

- מצב התחלתי < false, 0, Unknown, 9 >.

- מצב ביניים < *, *, *, 10-20 > חוץ ממצב סופי.

- מצב סופי > 20 , Unknown , * , * .

ככל שמרחיבים את שעות הפעילות של המרפאה עולה המספר הממוצע של חולים ששורדים:

