

## A - P2

<b>Student/An/Grupa:</b>									
<p><b>1. (1p)</b> Considerăm următorul grid în care un agent se poate deplasa și următoarele observații de tip <math>(s, a, s', R(s,a,s'))</math>: (B, East, C, 2), (C, South, E, 4), (C, East, A, D), (B, East, C, 2). Valorile inițiale ale funcției Q sunt 0. Presupunem <math>\alpha = 0.5</math>, <math>\gamma = 1</math>. Actualizați valorile Q. Care este politica recomandată pentru starea C?</p> <div><table><tr><td></td><td>A</td><td></td></tr><tr><td>B</td><td>C</td><td>D</td></tr><tr><td></td><td>E</td><td></td></tr></table></div>		A		B	C	D		E	
	A								
B	C	D							
	E								
<p><b>2. (0.75p)</b> Care este rolul parametrilor <math>\alpha</math>, <math>\gamma</math> și epsilon din cadrul metodei <i>epsilon-greedy Q-learning</i>? Care este efectul individual al setării fiecăruia la valoarea 0?</p>									
<p><b>3.</b> Considerăm rețeaua bayesiană de mai jos. Tabelele de probabilități condiționale sunt: <math>P(S) = 0.25</math>; <math>P(R   S) = 0.7</math>, <math>P(R   \neg S) = 0.3</math>; <math>P(Wo   R) = 0.7</math>, <math>P(Wo   \neg R) = 0.2</math>; <math>P(We   R, Wo) = 0.12</math>, <math>P(We   R, \neg Wo) = 0.25</math>, <math>P(We   \neg R, Wo) = 0.1</math>, <math>P(We   \neg R, \neg Wo) = 0.08</math>; <math>P(M   We) = 0.02</math>, <math>P(M   \neg We) = 0.42</math>; <math>P(B   S, Wo) = 0.8</math>, <math>P(B   S, \neg Wo) = 0.4</math>, <math>P(B   \neg S, Wo) = 0.4</math>, <math>P(B   \neg S, \neg Wo) = 0.4</math>.</p> <div></div>									

**a) (1p)** Calculați următoarele probabilități marginale: a) de a ploua b) de a fi râme afară c) de a fi păsări pe gazon.

**b) (0.5p)** Precizați cel puțin două relații de independență condițională din rețeaua bayesiană de mai sus. Justificați răspunsul.

**4.** Considerăm o problemă de planificare în care scopul este de a pune două baterii într-o lanternă. Operațiile posibile sunt: *PlaceCap()* pune capacul lanternei, *RemoveCap()* scoate capacul, *Insert(i)* inserează bateria în lanternă.

**a) (0.75p)** Descrieți cele trei operații posibile în limbajul STRIPS/ADL.

**b) (0.25p)** Considerați starea inițială în care lanterna are capacul pus. Starea finală este cea în care ambele baterii sunt puse în lanterna și capacul înșurubat. Descrieți cele două stări utilizând predicate.

**c) (0.75p)** Construiți un plan incomplet care să conțină cel puțin trei acțiuni, utilizând algoritmul de planificare cu ordine parțială. Specificați care sunt condițiile deschise și acțiunile selectate.