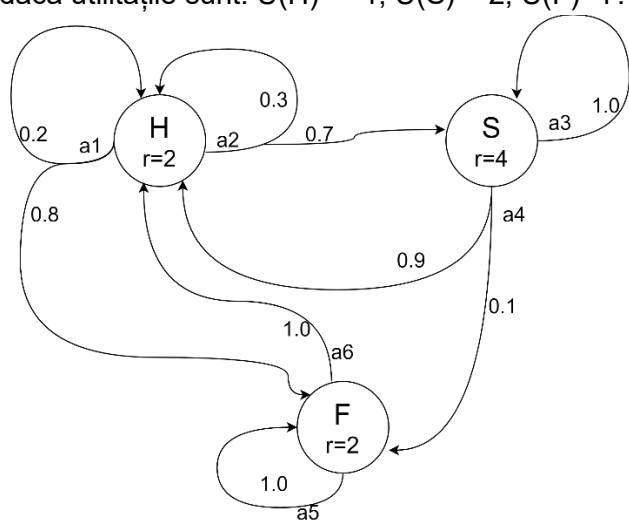


Important: Punctajul pentru exerciții se acordă doar pe baza unei rezolvări detaliate (scrieți mai întâi formulele și apoi înlocuiți cu valorile corespunzătoare)

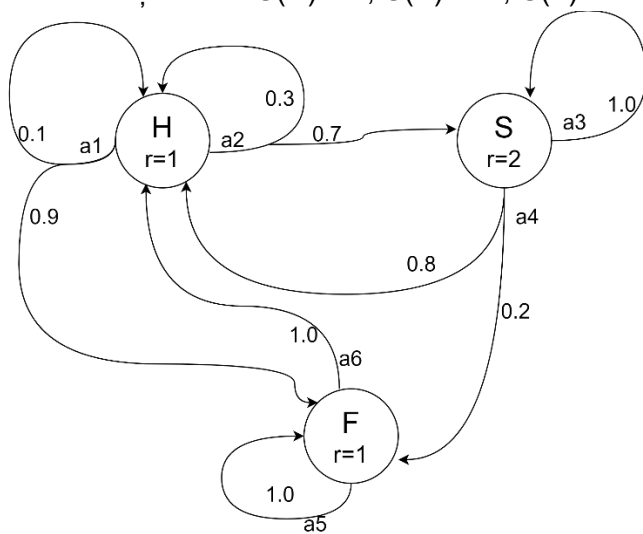
Subiect 1: Diagrama de mai jos modelează un proces de decizie Markov care surprinde locațiile în care te vei afla de sărbători și tranzițiile dintre acestea. Cele 3 stări sunt: H (acasă), S (la un magazin de cadouri) sau F (la un prieten acasă). Acțiunile sunt: a1 („mergi la un prieten acasă” – cu o mică probabilitate întâlnirea se poate amâna), a2 („cumpără cadou – online sau de la magazin”), a3 („rămâi la magazin”), a4 („mergi de la magazin acasă” – aici accidental te poți întâlni cu un prieten pe stradă și ajungi să mergi la el acasă), a5 („rămâi la prieten acasă”), a6 („mergi de la prieten, acasă la tine”). Recompensele primite în fiecare stare sunt marcate cu r . Conform principiului Maximum Expected Utility, ce acțiune va fi realizată din starea H, dacă utilitățile sunt: $U(H) = -1$, $U(S) = 2$, $U(F) = 1$?



Subiect 2: Presupunând că factorul de discount este 0.5 și că utilitățile inițiale sunt cele fixate la primul exercițiu, aplicați o iterație a algoritmului Value Iteration pentru a calcula utilitățile de la pasul următor.

Important: Punctajul pentru exerciții se acordă doar pe baza unei rezolvări detaliate (scrieți mai întâi formulele și apoi înlocuiți cu valorile corespunzătoare)

Subiect 1: Diagrama de mai jos modelează un proces de decizie Markov care surprinde locațiile în care te vei afla de sărbători și tranzițiile dintre acestea. Cele 3 stări sunt: H (acasă), S (la un magazin de cadouri) sau F (la un prieten acasă). Acțiunile sunt: a1 („mergi la un prieten acasă” – cu o mică probabilitate întâlnirea se poate amâna), a2 („cumpără cadou – online sau de la magazin”), a3 („rămâi la magazin”), a4 („mergi de la magazin acasă” – aici accidental te poți întâlni cu un prieten pe stradă și ajungi să mergi la el acasă), a5 („rămâi la prieten acasă”), a6 („mergi de la prieten, acasă la tine”). Recompensele primite în fiecare stare sunt marcate cu r. Conform principiului Maximum Expected Utility, ce acțiune va fi realizată din starea H, dacă utilitățile sunt: $U(H) = 2$, $U(S) = -1$, $U(F) = 1$?



Subiect 2: Presupunând că factorul de discount este 0.5 și că utilitățile inițiale sunt cele fixate la primul exercițiu, aplicați o iterație a algoritmului Value Iteration pentru a calcula utilitățile de la pasul următor.