

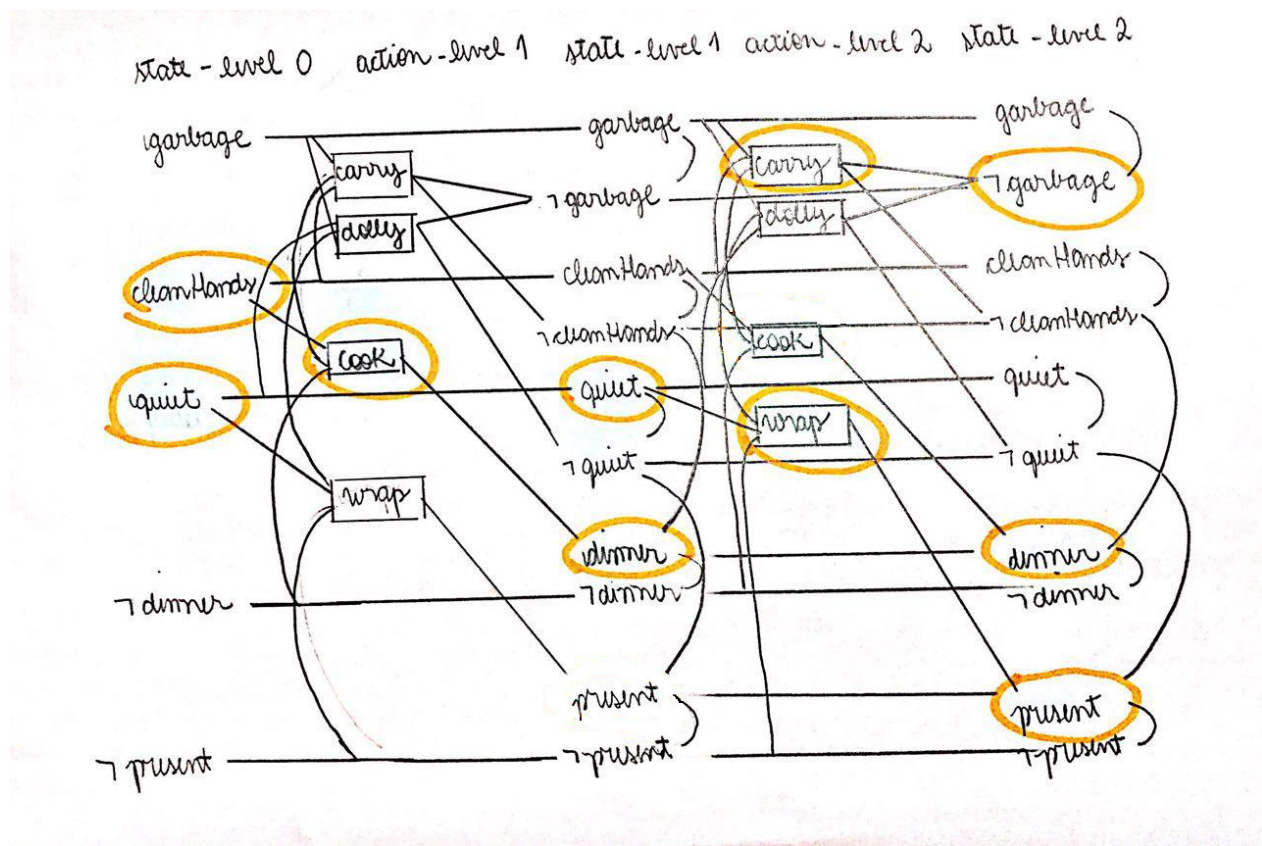
Nome: Luís Felipe de Melo Costa Silva
Número USP: 9297961

Lista de Exercícios 3 - MAC0425

Exercício 10.6

Quando os efeitos negativos de uma ação são descartados, nós adicionamos apenas os efeitos positivos, aumentando o conjunto de literais em que podemos encontrar a solução. Nesse caso, estamos simplificando as condições, tornando o problema menos restrito e mais geral, acabando com um problema mais fácil do que o original.

Exercício 10.9



- Começamos o Graphplan com os seguintes literais e os que não temos da meta no *state-level 0*.

- Então, aplicamos as ações possíveis a partir desses literais em *action-level 1*. Podemos observar que:
 - *carry* está em mutex com *cook* porque *carry* elimina a pré-condição *cleanHands* de *cook*. (Interferência)
 - *carry* está em mutex com a manutenção para garbage. (Efeitos inconsistentes)
 - *dolly* está em mutex com *wrap* porque *dolly* elimina a pré-condição *quiet* de *wrap*. (Interferência)
- Chegamos ao *state-level 1*, onde todas os literais são mutex uma com a outra por suporte inconsistente (todos possuem a negação no mesmo nível). Além disso, $\neg quiet$ está em mutex com *present* e $\neg cleanHands$ está em mutex com *dinner* pelo mesmo motivo. Temos dois possíveis conjuntos de soluções: [*carry, cook, wrap*] e [*dolly, cook, wrap*]. Nenhuma delas é válida porque *carry* está em mutex com *cook* e *dolly* está em mutex com *wrap*. Faremos mais uma expansão do grafo.
- Prosseguimos para o *action-level 2*, que é idêntico ao *action-level 1* e gera o *state-level 2*, semelhante ao *state-level 1*.

Para encontrarmos o plano para resolver esse problema, começamos selecionando os literais da meta em *state-level 2*. Então, pegamos ações no *action-level 2* que não estão em mutex entre si (*carry, wrap*) e as pré-condições dessas ações (*quiet*) e a ação de manutenção dos literais que não foram contemplados com as ações escolhidas (*dinner*), que estão no *state-level 1*. Olhamos para o *action-level 1* e procuramos as ações que geram *dinner* e *quiet*. A ação *cook* gera *dinner*, e *quiet* é gerado por uma ação de manutenção. Logo, nosso plano é [*cook, carry, wrap*]. Existem outros planos possíveis, por exemplo: [*wrap, cook, dolly*].

Exercício MDP

Exercício Q-Learning

Para esse exercício vamos usar a fórmula do *Q-value update*, que consiste em:

$$Q(s, a) \leftarrow (1 - \alpha) \cdot Q(s, a) + \alpha(r + \gamma \cdot \max_{a' \in A} Q(s', a'))$$

O enunciado nos pede para usar $\alpha = 1$ e $\gamma = 0,9$. Com esses valores, nossa fórmula ficará:

$$Q(s, a) \leftarrow (r + 0,9 \cdot \max_{a' \in A} Q(s', a'))$$

Começando com a tabela zerada e fazendo os cálculos para os episódios pedidos, teremos:

1.

$$Q(3, S) = -10 + 0,9 \cdot 0 = -10$$

$$Q(3, N) = 0 + 0,9 \cdot 0 = 0$$

$$Q(1, N) = -10 + 0,9 \cdot 0 = -10$$

$$Q(1, O) = 10 + 0,9 \cdot 0 = 10$$

2.

$$Q(2, L) = 0 + 0,9 \cdot 0 = 0$$

$$Q(3, N) = 0 + 0,9 \cdot 0 = 0$$

$$Q(1, N) = -10 + 0,9 \cdot 0 = -10$$

$$Q(1, L) = -10 + 0,9 \cdot 0 = -10$$

$$Q(1, O) = 10 + 0,9 \cdot 0 = 10$$

As tabelas finais Q dos episódios calculados estão na próxima página.

Table 1: Tabela do episódio da experiência 1

s/a	N	S	L	O
0	0	0	0	0
1	-10	0	0	10
2	0	0	0	0
3	0	-10	0	0

Table 2: Tabela do episódio da experiência 2

s/a	N	S	L	O
0	0	0	0	0
1	-10	0	-10	10
2	0	0	0	0
3	0	0	0	0