Un mdp è un framework per definire l’ambiente (stati, azioni, funzioni,ecc…) con il quale il nostro agente (IA o programmazione dinamica(algoritmo VI)) andra ad interagire per apprendere e risolvere il task.

E’ quadrupla S,A,P,R,gamma dove S è l’insieme dei possibili stati, ed ogni stato ha delle variabili che lo caratterizzano, e queste variabili possono assumere un valore all’interno di un certo dominio (anche semplicemente booleane) e altre quindi un vettore di tutti gli attacchi classificabili dall’intrucion detection system (ho gli attacchi nello stato tanti quanti ne conosce l’IDS e per ognuno dia ha una probabilità che indica se sono sotto attacco da quella tipologia di attacco).

A sono le azioni che possono essere intraprese all’istante t per andare su un nuovo stato t+1 con una certa probabilità P, con cui un’’azione venga scelta ed è data da una matrice StxAtxSt+1 e questa stessa matrice mi da o la ricompensa o il costo (asseconda del segno) che mi rappresenta la bonta di quell’azione in quello stato che mi portato sul nuovo (se l’azione ha senso reward alta, se non centra nulla bassa, quindi costo per segno negativo). Nel caso dei sistemi stazionari probabilità e reward non cambiano anche se cambia l’istante temporale, perche se mi trovo in St e vado in St+1 con sempre le stesse azioni, avro sempre stessa bonta. Cambia nel caso di sistemi non stazionari ovvero St puo variare al variare di t (non solo aggiungo nuova azione, rimuovo un’attacco, aggiungo variabile,ecc…)

L’obiettivo è di massimizzare la ricompensa, ed il comportamento dell’agente è descritto da una policy pigreco: SxA ovvero la probabilità con cui l’agente scelga l’azione At in St.

Infine in S ovvero gli stati trovo grazie a T, una funzione target, un sottoinsieme di S che sono gli stati di non attacco, stati target sicuri dove il mio agente si arresta, così riduco la dimensione. Infine ho anche una funzione PC ovvero pre condizioni che mi determina un sottoinsieme delle azioni applicabile in un determinato stato (se ho un certo attacco alcune contromisure sono inutili quindi non faccio determinate azioni se non mi aiutano a patchare; sto subendo sql injection ma vado a petchare cross site scripting, inutile).