Search Problem $\begin{cases} \text{DFS} \\ \text{BFS} \\ \text{UCS} \\ \text{Greedy Search} \\ A^* \text{ Search} \end{cases}$

Admissible: $0 \le h(n) \le h^*(n)$

$h^*(n)$ is the true cost to a nearest goal

Tree Search: can expand a state twice

Graph Search: never expand a state twice

Consistent: $h(A) - h(C) \le cost(A \text{ to } C)$

	Tree Search	Graph Search
admissable	optimal	X
consistent	optimal	optimal

Markov Decision Processes

The Bellman Equations

$$Q^*(s,a) = \sum_{s'} T(s,a,s')[R(s,a,s') + \gamma V^*(s')]$$

$$V^*(s) = \max_a Q^*(s,a)$$

Value Iteration

$$V_{k+1}(s) \leftarrow \max_a \sum_{s'} T(s,a,s')[R(s,a,s') + \gamma V_k(s')]$$

Complexity of each iteration: $O(S^2A)$

Policy Iteration

Evaluation: $O(S^2)$, do several pass until convergence

$$V_{k+1}^{\pi_i}(s) \leftarrow \sum_{s'} T(s,\pi_i(s),s')[R(s,\pi_i(s),s') + \gamma V_k^{\pi_i}(s')]$$

Improvement:

$$\pi_{i+1}(s) \leftarrow \arg\max_a \sum_{s'} T(s,a,s')[R(s,a,s') + \gamma V^{\pi_i}(s')]$$

If solved by a linear system, runtime is $O(S^3)$

Game

Alpha-Beta Pruning

α: MAX's best option on the path to root

β: MIN's best option on the path to root

```
    min-value
def max-value(state, α, β):
        +∞
    v = -∞
    for each successor of state:
              min
        v = max(v, value(successor, α, β))
              v≤α
        if  v≥β  return v
              min      β
        α = max(α, v)
    return v
```

Reinforcement Learning

Temporal Difference Learning

$$sample = R(s,\pi(s),s') + \gamma V^{\pi}(s')$$

$$V^{\pi}(s) = (1-\alpha)V^{\pi}(s) + \alpha \cdot sample$$

Q-Learning

$$sample = R(s,a,s') + \gamma \max_{a'} Q(s',a')$$

$$Q(s,a) = (1-\alpha)Q(s,a) + \alpha \cdot sample$$

$$\pi(s) = \arg\max_a Q(s,a)$$

Feature-Based Representations

$$V(s) = w_1 f_1(s) + w_2 f_2(s) + \cdots + w_n f_n(s)$$

$$Q(s,a) = w_1 f_1(s,a) + w_2 f_2(s,a) + \cdots + w_n f_n(s,a)$$

$$\left.\begin{array}{l} diff = [r - \gamma \max_{a'} Q(s',a')] - Q(s,a) \\ w_i = w_i + \alpha\,[diff] \cdot f_i(s,a) \end{array}\right\} \text{(for transition} = (s,a,r,s')\text{)}$$

ε-greedy exploration

probability ε: act randomly

probability $1-\varepsilon$: act on policy

CSP

Backtracking Search
= DFS + variable-ordering + fail-on-violation

Forward Checking
Cross off values that violate a constraint when added to the existing assignment

Arc Consistency: Delete from tail
An arc $X \to Y$ is consistent iff for every x in the tail there's some y in the head which could be assigned without violating a constraint.

Minimum Remaining Values
Least Constraining Value

Tree-Structured CSPs (No Loop)
Can be solved in $O(nd^2)$ time

Nearly Tree-Structured CSPs
Cutset Conditioning: $O(d^c (n-c) d^2)$

Logic
Entailment: $\alpha \models \beta$ (α entails β)
iff every world where α is true, β is also true
α-worlds are a subset of the β-worlds
(Truth table is very helpful)

successor-state axiom
$X_t \Leftrightarrow [X_{t-1} \wedge \neg(\text{some action}_{t-1} \text{ made it false})] \vee [\neg X_{t-1} \wedge (\text{some action}_{t-1} \text{ made it true})]$

Convert a sentence to CNF
① drop biconditionals
$\alpha \Leftrightarrow \beta \equiv (\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)$
② drop implications
$\alpha \Rightarrow \beta \equiv \neg\alpha \vee \beta$
③ move "not" inwards
$\neg(\alpha \vee \beta) \equiv (\neg\alpha \wedge \neg\beta)$
$\neg(\alpha \wedge \beta) \equiv (\neg\alpha \vee \neg\beta)$
④ move "or" inwards and "and" outwards
$\alpha \vee (\beta \wedge \gamma) \equiv (\alpha \vee \beta) \wedge (\alpha \vee \gamma)$

Efficient SAT solvers
DPLL
Pure literals: if all occurrences of a symbol in as-yet-unsatisfied clauses have the same sign, then give the symbol that value
$(\bar{A} \vee B) \wedge (\bar{A} \vee C) \wedge (D \vee \bar{E}) \wedge (\bar{B} \vee \bar{C})$
Unit clauses: if a clause is left with a single literal, set symbol to satisfy clause
$(A) \wedge (A \vee B) \wedge (C \vee D)$

Bayes net
$P(X_1, X_2 \cdots X_n) = \prod_i P(X_i \mid Parents(X_i))$

Every variable is conditionally independant of its non-descendants given its parents
A variable's Markov blanket consists of parents, children, children's other parents.
Every variable is conditionally independant of all other variables given its Markov blanket.

Variable Elimination

$P(B) \cdot P(E)\ P(A|B.E)\ P(j|A)\ P(m|A)$

a) Eliminate A

$$\left\{\begin{array}{l} P(A|B.E) \\ P(j|A) \\ P(m|A) \end{array}\right. \Rightarrow P(j.m|B.E)$$

$P(B)\ P(E)\ P(j.m|B.E)$

b) Eliminate E

$$\left\{\begin{array}{l} P(E) \\ P(j,m|B.E) \end{array}\right. \Rightarrow P(j.m|B)$$

Prior Sampling

For $i=1,2,\cdots,n$ (in topological order)

Sample X_i from $P(X_i | parents(X_i))$

Return $(X_1, X_2, \cdots, X_n)$

Rejection Sampling

If x_i not consistent with evidence

Reject

Likelihood Weighting

weight each sample by probability of evidence variables given parents

Gibbs Sampling

a) Fix evidence

b) Initialize other variables randomly

c) Repeat:

Choose an non-evidence variable X

Resample X from $P(X | markov\_blanket(X))$

Hidden Markov Models

Filtering: $O(|X|^2)$

$$P(X_t|e_{1:t-1}) = \sum_{x_{t-1}} P(X_t|x_{t-1}) P(x_{t-1}|e_{1:t-1})$$

$$P(X_t|e_{1:t}) \propto P(X_t|e_{1:t-1}) \cdot P(e_t|X_t)$$

Passage of time

$$\begin{aligned} P(X_t|e_{1:t}) &= P(X_t|e_{1:t-1}, e_t) \\ &= \alpha P(e_t|X_t, e_{1:t-1}) \cdot P(X_t|e_{1:t-1}) \\ &= \alpha P(e_t|X_t) \cdot P(X_t|e_{1:t-1}) \\ &= \alpha P(e_t|X_t) \cdot \sum_{x_{t-1}} P(X_t|x_{t-1}, e_{1:t-1}) \cdot P(x_{t-1}|e_{1:t-1}) \\ &= \alpha P(e_t|X_t) \sum_{x_{t-1}} P(X_t|x_{t-1}) \cdot P(x_{t-1}|e_{1:t-1}) \end{aligned}$$

Most likely Explanation: $\arg\max_{x_{1:t}} P(x_{1:t}|e_{1:t})$

Viterbi algorithm

$$\begin{aligned} m_t[X_t] &= \max_{x_{1:t-1}} P(x_{1:t}, e_{1:t}) \\ &= P(e_t|X_t) \max_{x_{t-1}} P(X_t|x_{t-1}) m_{t-1}[x_{t-1}] \end{aligned}$$

Time complexity: $O(|X|^2 \cdot T)$

Space complexity: $O(|X| \cdot T)$

Particle Filtering

a) Elapse Time: Each particle is moved by sampling its next position from the transition model

b) Observe: down weight samples based on the evidence

$w(x) = P(e|x)$

c) Resample particles according to their weights

No.

Date

Value of perfect Information

$VPI(E'|e) = (\sum_{e'} P(e'|e)\, MEU(e,e') - MEU(e))$

a) $\forall E', e : VPI(E'|e) \geq 0$

b) $VPI(E_j, E_k|e) \neq VPI(E_j|e) + VPI(E_k|e)$

c) $VPI(E_j, E_k|e) = VPI(E_j|e) + VPI(E_k|e, E_j)$

$= VPI(E_k|e) + VPI(E_j|e, E_k)$

Naive Bayes

$P(Y, F_1, F_2 \cdots F_n) = P(Y) \prod_i P(F_i|Y)$

Maximum Likelihood

Relative frequencies are the maximum likelihood estimates

$P_{ML}(x) = \frac{count(x)}{total\ samples}$

Laplace Smoothing

$P_{LAP,k}(x) = \frac{count(x) + k}{N + k|X|}$

Entropy

$H(p_1, p_2 \cdots p_n) = E\left[\log_2 \frac{1}{p_i}\right]$

$= \sum_{i=1}^{n} -p_i \log_2 p_i$

HMM: Stationary Distributions

e.g. $\begin{cases} P_\infty(sun) = P(sun|sun) P_\infty(sun) + P(sun|rain) P_\infty(rain) \\ P_\infty(rain) = P(rain|sun) P_\infty(sun) + P(rain|rain) P_\infty(rain) \\ P_\infty(sun) + P_\infty(rain) = 1 \end{cases}$