CHAPTER 1

INTRODUCTION

CHAPTER 2

INTELLIGENT AGENTS

function TABLE-DRIVEN-AGENT($percept$) **returns** an action
  persistent: $percepts$, a sequence, initially empty
    $table$, a table of actions, indexed by percept sequences, initially fully specified

  append $percept$ to the end of $percepts$
  $action \leftarrow$ LOOKUP($percepts$, $table$)
  return $action$

Figure 2.7 The TABLE-DRIVEN-AGENT program is invoked for each new percept and returns an action each time. It retains the complete percept sequence in memory.

function REFLEX-VACUUM-AGENT([$location$,$status$]) **returns** an action

  if $status = Dirty$ **then return** $Suck$
  else if $location = A$ **then return** $Right$
  else if $location = B$ **then return** $Left$

Figure 2.8 The agent program for a simple reflex agent in the two-location vacuum environment. This program implements the agent function tabulated in Figure **??**.

function SIMPLE-REFLEX-AGENT($percept$) **returns** an action
  persistent: $rules$, a set of condition–action rules

  $state \leftarrow$ INTERPRET-INPUT($percept$)
  $rule \leftarrow$ RULE-MATCH($state, rules$)
  $action \leftarrow rule$.ACTION
  return $action$

Figure 2.10 A simple reflex agent. It acts according to a rule whose condition matches the current state, as defined by the percept.

function MODEL-BASED-REFLEX-AGENT($percept$) **returns** an action
  persistent: $state$, the agent's current conception of the world state
    $transition\_model$, a description of how the next state depends on
      the current state and action
    $sensor\_model$, a description of how the current world state is reflected
      in the agent's percepts
    $rules$, a set of condition–action rules
    $action$, the most recent action, initially none

  $state \leftarrow$ UPDATE-STATE($state, action, percept, transition\_model, sensor\_model$)
  $rule \leftarrow$ RULE-MATCH($state, rules$)
  $action \leftarrow rule$.ACTION
  return $action$

Figure 2.12 A model-based reflex agent. It keeps track of the current state of the world, using an internal model. It then chooses an action in the same way as the reflex agent.

CHAPTER 3

SOLVING PROBLEMS BY SEARCHING

```
function BEST-FIRST-SEARCH(problem, f) returns a solution node or failure
  node ← NODE(STATE=problem.INITIAL)
  frontier ← a priority queue ordered by f, with node as an element
  reached ← a lookup table, with one entry with key problem.INITIAL and value node
  while not IS-EMPTY(frontier) do
    node ← POP(frontier)
    if problem.IS-GOAL(node.STATE) then return node
    for each child in EXPAND(problem, node) do
      s ← child.STATE
      if s is not in reached or child.PATH-COST < reached[s].PATH-COST then
        reached[s] ← child
        add child to frontier
  return failure

function EXPAND(problem, node) yields nodes
  s ← node.STATE
  for each action in problem.ACTIONS(s) do
    s′ ← problem.RESULT(s, action)
    cost ← node.PATH-COST + problem.ACTION-COST(s, action, s′)
    yield NODE(STATE=s′, PARENT=node, ACTION=action, PATH-COST=cost)
```

Figure 3.7 The best-first search algorithm, and the function for expanding a node. The data structures used here are described in Section **??**. See Appendix B for **yield**.

```
function BREADTH-FIRST-SEARCH(problem) returns a solution node or failure
  node ← NODE(problem.INITIAL)
  if problem.IS-GOAL(node.STATE) then return node
  frontier ← a FIFO queue, with node as an element
  reached ← {problem.INITIAL}
  while not IS-EMPTY(frontier) do
    node ← POP(frontier)
    for each child in EXPAND(problem, node) do
      s ← child.STATE
      if problem.IS-GOAL(s) then return child
      if s is not in reached then
        add s to reached
        add child to frontier
  return failure

function UNIFORM-COST-SEARCH(problem) returns a solution node, or failure
  return BEST-FIRST-SEARCH(problem, PATH-COST)
```

Figure 3.9 Breadth-first search and uniform-cost search algorithms.

```
function ITERATIVE-DEEPENING-SEARCH(problem) returns a solution node or failure
  for depth = 0 to ∞ do
    result ← DEPTH-LIMITED-SEARCH(problem, depth)
    if result ≠ cutoff then return result

function DEPTH-LIMITED-SEARCH(problem, ℓ) returns a node or failure or cutoff
  frontier ← a LIFO queue (stack) with NODE(problem.INITIAL) as an element
  result ← failure
  while not IS-EMPTY(frontier) do
    node ← POP(frontier)
    if problem.IS-GOAL(node.STATE) then return node
    if DEPTH(node) > ℓ then
      result ← cutoff
    else if not IS-CYCLE(node) do
      for each child in EXPAND(problem, node) do
        add child to frontier
  return result
```

Figure 3.12 Iterative deepening and depth-limited tree-like search. Iterative deepening repeatedly applies depth-limited search with increasing limits. It returns one of three different types of values: either a solution node; or *failure*, when it has exhausted all nodes and proved there is no solution at any depth; or *cutoff*, to mean there might be a solution at a deeper depth than ℓ. This is a tree-like search algorithm that does not keep track of *reached* states, and thus uses much less memory than best-first search, but runs the risk of visiting the same state multiple times on different paths. Also, if the IS-CYCLE check does not check *all* cycles, then the algorithm may get caught in a loop.

```
function BIBF-SEARCH(problem_F, f_F, problem_B, f_B) returns a solution node, or failure
  node_F ← NODE(problem_F.INITIAL)                          // Node for a start state
  node_B ← NODE(problem_B.INITIAL)                          // Node for a goal state
  frontier_F ← a priority queue ordered by f_F, with node_F as an element
  frontier_B ← a priority queue ordered by f_B, with node_B as an element
  reached_F ← a lookup table, with one key node_F.STATE and value node_F
  reached_B ← a lookup table, with one key node_B.STATE and value node_B
  solution ← failure
  while not TERMINATED(solution, frontier_F, frontier_B) do
    if f_F(TOP(frontier_F)) < f_B(TOP(frontier_B)) then
      solution ← PROCEED(F, problem_F frontier_F, reached_F, reached_B, solution)
    else solution ← PROCEED(B, problem_B, frontier_B, reached_B, reached_F, solution)
  return solution

function PROCEED(dir, problem, frontier, reached, reached_2, solution) returns a solution
        // Expand node on frontier; check against the other frontier in reached_2.
        // The variable "dir" is the direction: either F for forward or B for backward.
  node ← POP(frontier)
  for each child in EXPAND(problem, node) do
    s ← child.STATE
    if s not in reached or PATH-COST(child) < PATH-COST(reached[s]) then
      reached[s] ← child
      add child to frontier
      if s is in reached_2 then
        solution_2 ← JOIN-NODES(dir, child, reached_2[s]))
        if PATH-COST(solution_2) < PATH-COST(solution) then
          solution ← solution_2
  return solution
```

Figure 3.14 Bidirectional best-first search keeps two frontiers and two tables of reached states. When a path in one frontier reaches a state that was also reached in the other half of the search, the two paths are joined (by the function JOIN-NODES) to form a solution. The first solution we get is not guaranteed to be the best; the function TERMINATED determines when to stop looking for new solutions.

```
function RECURSIVE-BEST-FIRST-SEARCH(problem) returns a solution or failure
    solution, fvalue ← RBFS(problem, NODE(problem.INITIAL), ∞)
  return solution

function RBFS(problem, node, f_limit) returns a solution or failure, and a new f-cost limit
   if problem.IS-GOAL(node.STATE) then return node
   successors ← LIST(EXPAND(node))
   if successors is empty then return failure, ∞
   for each s in  successors do        // update f with value from previous search
      s.f ← max(s.PATH-COST + h(s), node.f))
   while true do
      best ← the node in successors with lowest f-value
      if best.f > f_limit then return failure, best.f
      alternative ← the second-lowest f-value among successors
      result, best.f ← RBFS(problem, best, min(f_limit, alternative))
      if result ≠ failure then return result, best.f
```

Figure 3.22 The algorithm for recursive best-first search.

CHAPTER 4

SEARCH IN COMPLEX ENVIRONMENTS

function HILL-CLIMBING($problem$) **returns** a state that is a local maximum
  $current \leftarrow problem$.INITIAL
  while $true$ **do**
    $neighbor \leftarrow$ a highest-valued successor state of $current$
    if VALUE($neighbor$) $\leq$ VALUE($current$) **then return** $current$
    $current \leftarrow neighbor$

Figure 4.2 The hill-climbing search algorithm, which is the most basic local search technique. At each step the current node is replaced by the best neighbor.

function SIMULATED-ANNEALING($problem$, $schedule$) **returns** a solution state
  $current \leftarrow problem$.INITIAL
  for t = 1 **to** ∞ **do**
    $T \leftarrow schedule(t)$
    if T = 0 **then return** $current$
    $next \leftarrow$ a randomly selected successor of $current$
    $\Delta E \leftarrow$ VALUE($current$) – VALUE($next$)
    if $\Delta E > 0$ **then** $current \leftarrow next$
    else $current \leftarrow next$ only with probability $e^{-\Delta E/T}$

Figure 4.4 The simulated annealing algorithm, a version of stochastic hill climbing where some downhill moves are allowed. The $schedule$ input determines the value of the “temperature” T as a function of time.

```
function GENETIC-ALGORITHM(population, fitness) returns an individual
  repeat
      weights ← WEIGHTED-BY(population, fitness)
      population2 ← empty list
      for i = 1 to SIZE(population) do
          parent1, parent2 ← WEIGHTED-RANDOM-CHOICES(population, weights, 2)
          child ← REPRODUCE(parent1, parent2)
          if (small random probability) then child ← MUTATE(child)
          add child to population2
      population ← population2
  until some individual is fit enough, or enough time has elapsed
  return the best individual in population, according to fitness

function REPRODUCE(parent1, parent2) returns an individual
  n ← LENGTH(parent1)
  c ← random number from 1 to n
  return APPEND(SUBSTRING(parent1, 1, c), SUBSTRING(parent2, c + 1, n))
```

Figure 4.7 A genetic algorithm. Within the function, *population* is an ordered list of individuals, *weights* is a list of corresponding fitness values for each individual, and *fitness* is a function to compute these values.

```
function AND-OR-SEARCH(problem) returns a conditional plan, or failure
  return OR-SEARCH(problem, problem.INITIAL, [ ])

function OR-SEARCH(problem, state, path) returns a conditional plan, or failure
  if problem.IS-GOAL(state) then return the empty plan
  if IS-CYCLE(path) then return failure
  for each action in problem.ACTIONS(state) do
      plan ← AND-SEARCH(problem, RESULTS(state, action), [state] + path])
      if plan ≠ failure then return [action] + plan]
  return failure

function AND-SEARCH(problem, states, path) returns a conditional plan, or failure
  for each s_i in states do
      plan_i ← OR-SEARCH(problem, s_i, path)
      if plan_i = failure then return failure
  return [if s_1 then plan_1 else if s_2 then plan_2 else ... if s_{n-1} then plan_{n-1} else plan_n]
```

Figure 4.10 An algorithm for searching AND–OR graphs generated by nondeterministic environments. A solution is a conditional plan that considers every nondeterministic outcome and makes a plan for each one.

function ONLINE-DFS-AGENT($problem$, s') **returns** an action
    s, a, the previous state and action, initially null
  persistent: $result$, a table mapping $(s,\ a)$ to s', initially empty
    $untried$, a table mapping s to a list of untried actions
    $unbacktracked$, a table mapping s to a list of states never backtracked to

  if $problem$.IS-GOAL(s') **then return** $stop$
  if s' is a new state (not in $untried$) **then** $untried[s'] \leftarrow problem$.ACTIONS($s'$)
  if s is not null **then**
    $result[s, a] \leftarrow s'$
    add s to the front of $unbacktracked[s']$
  if $untried[s']$ is empty **then**
    if $unbacktracked[s']$ is empty **then return** $stop$
    else $a \leftarrow$ an action b such that $result[s', b]$ = POP($unbacktracked[s']$)
  else $a \leftarrow$ POP($untried[s']$)
  $s \leftarrow s'$
  return a

Figure 4.20 An online search agent that uses depth-first exploration. The agent can safely explore only in state spaces in which every action can be "undone" by some other action.

function LRTA*-AGENT($problem$, s', h) **returns** an action
    s, a, the previous state and action, initially null
  persistent: $result$, a table mapping $(s,\ a)$ to s', initially empty
    H, a table mapping s to a cost estimate, initially empty

  if IS-GOAL(s') **then return** $stop$
  if s' is a new state (not in H) **then** $H[s'] \leftarrow h(s')$
  if s is not null **then**
    $result[s, a] \leftarrow s'$
    $H[s] \leftarrow \min_{b \in \text{ACTIONS}(s)} \text{LRTA*-COST}(s, b, result[s, b], H)$
  $a \leftarrow \operatorname{argmin}_{b \in \text{ACTIONS}(s)} \text{LRTA*-COST}(problem, s', b, result[s', b], H)$
  $s \leftarrow s'$
  return a

function LRTA*-COST($problem, s, a, s', H$) **returns** a cost estimate
  if s' is undefined **then return** $h(s)$
  else return $problem$.ACTION-COST(s, a, s') $+\ H[s']$

Figure 4.23 LRTA*-AGENT selects an action according to the values of neighboring states, which are updated as the agent moves about the state space.

CHAPTER 5

ADVERSARIAL SEARCH AND GAMES

```
function MINIMAX-SEARCH(game, state) returns an action
   player ← game.TO-MOVE(state)
   value, move ← MAX-VALUE(game, state)
   return move

function MAX-VALUE(game, state) returns a (utility, move) pair
   if game.IS-TERMINAL(state) then return game.UTILITY(state, player), null
   v ← −∞
   for each a in game.ACTIONS(state) do
      v2, a2 ← MIN-VALUE(game, game.RESULT(state, a))
      if v2 > v then
         v, move ← v2, a
   return v, move

function MIN-VALUE(game, state) returns a (utility, move) pair
   if game.IS-TERMINAL(state) then return game.UTILITY(state, player), null
   v ← +∞
   for each a in game.ACTIONS(state) do
      v2, a2 ← MAX-VALUE(game, game.RESULT(state, a))
      if v2 < v then
         v, move ← v2, a
   return v, move
```

Figure 5.3 An algorithm for calculating the optimal move using minimax—the move that leads to a terminal state with maximum utility, under the assumption that the opponent plays to minimize utility. The functions MAX-VALUE and MIN-VALUE go through the whole game tree, all the way to the leaves, to determine the backed-up value of a state and the move to get there.

```
function ALPHA-BETA-SEARCH(game, state) returns an action
   player ← game.TO-MOVE(state)
   value, move ← MAX-VALUE(game, state, −∞, +∞)
   return move

function MAX-VALUE(game, state, α, β) returns a (utility, move) pair
   if game.IS-TERMINAL(state) then return game.UTILITY(state, player), null
   v ← −∞
   for each a in game.ACTIONS(state) do
      v2, a2 ← MIN-VALUE(game, game.RESULT(state, a), α, β)
      if v2 > v then
         v, move ← v2, a
         α ← MAX(α, v)
      if v ≥ β then return v, move
   return v, move

function MIN-VALUE(game, state, α, β) returns a (utility, move) pair
   if game.IS-TERMINAL(state) then return game.UTILITY(state, player), null
   v ← +∞
   for each a in game.ACTIONS(state) do
      v2, a2 ← MAX-VALUE(game, game.RESULT(state, a), α, β)
      if v2 < v then
         v, move ← v2, a
         β ← MIN(β, v)
      if v ≤ α then return v, move
   return v, move
```

Figure 5.7 The alpha–beta search algorithm. Notice that these functions are the same as the MINIMAX-SEARCH functions in Figure **??**, except that we maintain bounds in the variables α and β, and use them to cut off search when a value is outside the bounds.

```
function MONTE-CARLO-TREE-SEARCH(state) returns an action
   tree ← NODE(state)
   while IS-TIME-REMAINING() do
      leaf ← SELECT(tree)
      child ← EXPAND(leaf)
      result ← SIMULATE(child)
      BACK-PROPAGATE(result, child)
   return the move in ACTIONS(state) whose node has highest number of playouts
```

Figure 5.11 The Monte Carlo tree search algorithm. A game tree, *tree*, is initialized, and then we repeat a cycle of SELECT / EXPAND / SIMULATE / BACK-PROPAGATE until we run out of time, and return the move that led to the node with the highest number of playouts.

CHAPTER 6

CONSTRAINT SATISFACTION PROBLEMS

```
function AC-3( csp) returns false if an inconsistency is found and true otherwise
  queue ← a queue of arcs, initially all the arcs in csp

  while queue is not empty do
    (X_i, X_j) ← POP(queue)
    if REVISE(csp, X_i, X_j) then
      if size of D_i = 0 then return false
      for each X_k in X_i.NEIGHBORS - {X_j} do
        add (X_k, X_i) to queue
  return true

function REVISE( csp, X_i, X_j) returns true iff we revise the domain of X_i
  revised ← false
  for each x in D_i do
    if no value y in D_j allows (x,y) to satisfy the constraint between X_i and X_j then
      delete x from D_i
      revised ← true
  return revised
```

Figure 6.3 The arc-consistency algorithm AC-3. After applying AC-3, either every arc is arc-consistent, or some variable has an empty domain, indicating that the CSP cannot be solved. The name "AC-3" was used by the algorithm's inventor (?) because it was the third version developed in the paper.

```
function BACKTRACKING-SEARCH(csp) returns a solution or failure
  return BACKTRACK(csp, { })

function BACKTRACK(csp, assignment) returns a solution or failure
  if assignment is complete then return assignment
  var ← SELECT-UNASSIGNED-VARIABLE(csp, assignment)
  for each value in ORDER-DOMAIN-VALUES(csp, var, assignment) do
     if value is consistent with assignment then
        add {var = value} to assignment
        inferences ← INFERENCE(csp, var, assignment)
        if inferences ≠ failure then
           add inferences to csp
           result ← BACKTRACK(csp, assignment)
           if result ≠ failure then return result
           remove inferences from csp
        remove {var = value} from assignment
  return failure
```

Figure 6.5 A simple backtracking algorithm for constraint satisfaction problems. The algorithm is modeled on the recursive depth-first search of Chapter **??**. The functions SELECT-UNASSIGNED-VARIABLE and ORDER-DOMAIN-VALUES, implement the general-purpose heuristics discussed in Section **??**. The INFERENCE function can optionally impose arc-, path-, or k-consistency, as desired. If a value choice leads to failure (noticed either by INFERENCE or by BACKTRACK), then value assignments (including those made by INFERENCE) are retracted and a new value is tried.

```
function MIN-CONFLICTS(csp, max_steps) returns a solution or failure
  inputs: csp, a constraint satisfaction problem
          max_steps, the number of steps allowed before giving up

  current ← an initial complete assignment for csp
  for i = 1 to max_steps do
     if current is a solution for csp then return current
     var ← a randomly chosen conflicted variable from csp.VARIABLES
     value ← the value v for var that minimizes CONFLICTS(csp, var, v, current)
     set var = value in current
  return failure
```

Figure 6.9 The MIN-CONFLICTS local search algorithm for CSPs. The initial state may be chosen randomly or by a greedy assignment process that chooses a minimal-conflict value for each variable in turn. The CONFLICTS function counts the number of constraints violated by a particular value, given the rest of the current assignment.

```
function TREE-CSP-SOLVER(csp) returns a solution, or failure
  inputs: csp, a CSP with components X, D, C

  n ← number of variables in X
  assignment ← an empty assignment
  root ← any variable in X
  X ← TOPOLOGICALSORT(X, root)
  for j = n down to 2 do
    MAKE-ARC-CONSISTENT(PARENT(X_j), X_j)
    if it cannot be made consistent then return failure
  for i = 1 to n do
    assignment[X_i] ← any consistent value from D_i
    if there is no consistent value then return failure
  return assignment
```

Figure 6.11 The TREE-CSP-SOLVER algorithm for solving tree-structured CSPs. If the CSP has a solution, we will find it in linear time; if not, we will detect a contradiction.

CHAPTER 7

LOGICAL AGENTS

function KB-AGENT($percept$) **returns** an $action$
  persistent: KB, a knowledge base
    t, a counter, initially 0, indicating time

  TELL(KB, MAKE-PERCEPT-SENTENCE($percept, t$))
  $action \leftarrow$ ASK(KB, MAKE-ACTION-QUERY(t))
  TELL(KB, MAKE-ACTION-SENTENCE($action, t$))
  $t \leftarrow t + 1$
  return $action$

Figure 7.1 A generic knowledge-based agent. Given a percept, the agent adds the percept to its knowledge base, asks the knowledge base for the best action, and tells the knowledge base that it has in fact taken that action.

```
function TT-ENTAILS?(KB, α) returns true or false
  inputs: KB, the knowledge base, a sentence in propositional logic
          α, the query, a sentence in propositional logic

  symbols ← a list of the proposition symbols in KB and α
  return TT-CHECK-ALL(KB, α, symbols, { })

function TT-CHECK-ALL(KB, α, symbols, model) returns true or false
  if EMPTY?(symbols) then
     if PL-TRUE?(KB, model) then return PL-TRUE?(α, model)
     else return true       //  when KB is false, always return true
  else
     P ← FIRST(symbols)
     rest ← REST(symbols)
     return (TT-CHECK-ALL(KB, α, rest, model ∪ {P = true})
            and
            TT-CHECK-ALL(KB, α, rest, model ∪ {P = false }))
```

Figure 7.10 A truth-table enumeration algorithm for deciding propositional entailment. (TT stands for truth table.) PL-TRUE? returns *true* if a sentence holds within a model. The variable *model* represents a partial model—an assignment to some of the symbols. The keyword **and** here is an infix function symbol in the pseudocode programming language, not an operator in proposition logic; it takes two arguments and returns *true* or *false*.

```
function PL-RESOLUTION(KB, α) returns true or false
  inputs: KB, the knowledge base, a sentence in propositional logic
          α, the query, a sentence in propositional logic

  clauses ← the set of clauses in the CNF representation of KB ∧ ¬α
  new ← { }
  while true  do
     for each pair of clauses C_i, C_j in clauses do
        resolvents ← PL-RESOLVE(C_i, C_j)
        if resolvents contains the empty clause then return true
        new ← new ∪ resolvents
     if new ⊆ clauses then return false
     clauses ← clauses ∪ new
```

Figure 7.13 A simple resolution algorithm for propositional logic. PL-RESOLVE returns the set of all possible clauses obtained by resolving its two inputs.

```
function PL-FC-ENTAILS?(KB, q) returns true or false
  inputs: KB, the knowledge base, a set of propositional definite clauses
          q, the query, a proposition symbol
  count ← a table, where count[c] is initially the number of symbols in clause c's premise
  inferred ← a table, where inferred[s] is initially false for all symbols
  queue ← a queue of symbols, initially symbols known to be true in KB

  while queue is not empty do
     p ← POP(queue)
     if p = q then return true
     if inferred[p] = false then
        inferred[p] ← true
        for each clause c in KB where p is in c.PREMISE do
           decrement count[c]
           if count[c] = 0 then add c.CONCLUSION to queue
  return false
```

Figure 7.15 The forward-chaining algorithm for propositional logic. The *agenda* keeps track of symbols known to be true but not yet "processed." The *count* table keeps track of how many premises of each implication are not yet proven. Whenever a new symbol p from the agenda is processed, the count is reduced by one for each implication in whose premise p appears (easily identified in constant time with appropriate indexing.) If a count reaches zero, all the premises of the implication are known, so its conclusion can be added to the agenda. Finally, we need to keep track of which symbols have been processed; a symbol that is already in the set of inferred symbols need not be added to the agenda again. This avoids redundant work and prevents loops caused by implications such as $P \Rightarrow Q$ and $Q \Rightarrow P$.

```
function DPLL-SATISFIABLE?(s) returns true or false
  inputs: s, a sentence in propositional logic

  clauses ← the set of clauses in the CNF representation of s
  symbols ← a list of the proposition symbols in s
  return DPLL(clauses, symbols, { })

function DPLL(clauses, symbols, model) returns true or false

  if every clause in clauses is true in model then return true
  if some clause in clauses is false in model then return false
  P, value ← FIND-PURE-SYMBOL(symbols, clauses, model)
  if P is non-null then return DPLL(clauses, symbols – P, model ∪ {P=value})
  P, value ← FIND-UNIT-CLAUSE(clauses, model)
  if P is non-null then return DPLL(clauses, symbols – P, model ∪ {P=value})
  P ← FIRST(symbols); rest ← REST(symbols)
  return DPLL(clauses, rest, model ∪ {P=true}) or
         DPLL(clauses, rest, model ∪ {P=false}))
```

Figure 7.17 The DPLL algorithm for checking satisfiability of a sentence in propositional logic. The ideas behind FIND-PURE-SYMBOL and FIND-UNIT-CLAUSE are described in the text; each returns a symbol (or null) and the truth value to assign to that symbol. Like TT-ENTAILS?, DPLL operates over partial models.

```
function WALKSAT(clauses, p, max_flips) returns a satisfying model or failure
  inputs: clauses, a set of clauses in propositional logic
          p, the probability of choosing to do a "random walk" move, typically around 0.5
          max_flips, number of value flips allowed before giving up

  model ← a random assignment of true/false to the symbols in clauses
  for each i = 1 to max_flips do
     if model satisfies clauses then return model
     clause ← a randomly selected clause from clauses that is false in model
     if RANDOM(0, 1) ≤ p then
        flip the value in model of a randomly selected symbol from clause
     else flip whichever symbol in clause maximizes the number of satisfied clauses
  return failure
```

Figure 7.18 The WALKSAT algorithm for checking satisfiability by randomly flipping the values of variables. Many versions of the algorithm exist.

```
function HYBRID-WUMPUS-AGENT(percept) returns an action
  inputs: percept, a list, [stench,breeze,glitter,bump,scream]
  persistent: KB, a knowledge base, initially the atemporal "wumpus physics"
              t, a counter, initially 0, indicating time
              plan, an action sequence, initially empty

  TELL(KB, MAKE-PERCEPT-SENTENCE(percept, t))
  TELL the KB the temporal "physics" sentences for time t
  safe ← {[x, y] : ASK(KB, OK^t_{x,y}) = true}
  if ASK(KB, Glitter^t) = true then
    plan ← [Grab] + PLAN-ROUTE(current, {[1,1]}, safe) + [Climb]
  if plan is empty then
    unvisited ← {[x, y] : ASK(KB, L^{t'}_{x,y}) = false for all t' ≤ t}
    plan ← PLAN-ROUTE(current, unvisited ∩ safe, safe)
  if plan is empty and ASK(KB, HaveArrow^t) = true then
    possible_wumpus ← {[x, y] : ASK(KB, ¬ W_{x,y}) = false}
    plan ← PLAN-SHOT(current, possible_wumpus, safe)
  if plan is empty then          // no choice but to take a risk
    not_unsafe ← {[x, y] : ASK(KB, ¬ OK^t_{x,y}) = false}
    plan ← PLAN-ROUTE(current, unvisited ∩ not_unsafe, safe)
  if plan is empty then
    plan ← PLAN-ROUTE(current, {[1, 1]}, safe) + [Climb]
  action ← POP(plan)
  TELL(KB, MAKE-ACTION-SENTENCE(action, t))
  t ← t + 1
  return action

function PLAN-ROUTE(current,goals,allowed) returns an action sequence
  inputs: current, the agent's current position
          goals, a set of squares; try to plan a route to one of them
          allowed, a set of squares that can form part of the route

  problem ← ROUTE-PROBLEM(current, goals,allowed)
  return SEARCH(problem)         // Any search algorithm from Chapter ??
```

Figure 7.20 A hybrid agent program for the wumpus world. It uses a propositional knowledge base to infer the state of the world, and a combination of problem-solving search and domain-specific code to choose actions. Each time HYBRID-WUMPUS-AGENT is called, it adds the percept to the knowledge base, and then either relies on a previously-defined plan or creates a new plan, and pops off the first step of the plan as the action to do next.

```
function SATPLAN(init, transition, goal, T_max) returns solution or failure
   inputs: init, transition, goal, constitute a description of the problem
           T_max, an upper limit for plan length

   for t = 0 to T_max do
      cnf ← TRANSLATE-TO-SAT(init, transition, goal, t)
      model ← SAT-SOLVER(cnf)
      if model is not null then
         return EXTRACT-SOLUTION(model)
   return failure
```

Figure 7.22 The SATPLAN algorithm. The planning problem is translated into a CNF sentence in which the goal is asserted to hold at a fixed time step t and axioms are included for each time step up to t. If the satisfiability algorithm finds a model, then a plan is extracted by looking at those proposition symbols that refer to actions and are assigned *true* in the model. If no model exists, then the process is repeated with the goal moved one step later.

CHAPTER 8

FIRST-ORDER LOGIC

CHAPTER 9

INFERENCE IN FIRST-ORDER LOGIC

function UNIFY(x, y, θ=*empty*) **returns** a substitution to make x and y identical, or *failure*
  if θ = *failure* **then return** *failure*
  else if $x = y$ **then return** θ
  else if VARIABLE?(x) **then return** UNIFY-VAR(x, y, θ)
  else if VARIABLE?(y) **then return** UNIFY-VAR(y, x, θ)
  else if COMPOUND?(x) **and** COMPOUND?(y) **then**
    return UNIFY(ARGS(x), ARGS(y), UNIFY(OP(x), OP(y), θ))
  else if LIST?(x) **and** LIST?(y) **then**
    return UNIFY(REST(x), REST(y), UNIFY(FIRST(x), FIRST(y), θ))
  else return *failure*

function UNIFY-VAR(var, x, θ) **returns** a substitution
  if $\{var/val\} \in \theta$ for some val **then return** UNIFY(val, x, θ)
  else if $\{x/val\} \in \theta$ for some val **then return** UNIFY(var, val, θ)
  else if OCCUR-CHECK?(var, x) **then return** *failure*
  else return add $\{var/x\}$ to θ

Figure 9.1 The unification algorithm. The arguments x and y can be any expression: a constant or variable, or a compound expression such as a complex sentence or term, or a list of expressions. The argument θ is a substitution, initially the empty substitution, but with $\{var/val\}$ pairs added to it as we recurse through the inputs, comparing the expressions element by element. In a compound expression such as $F(A, B)$, OP(x) field picks out the function symbol F and ARGS(x) field picks out the argument list (A, B).

function FOL-FC-ASK(KB, α) **returns** a substitution or *false*
  inputs: KB, the knowledge base, a set of first-order definite clauses
    α, the query, an atomic sentence

  while *true* **do**
    $new \leftarrow \{\}$ *// The set of new sentences inferred on each iteration*
    for each *rule* **in** KB **do**
      $(p_1 \wedge \ldots \wedge p_n \Rightarrow q) \leftarrow$ STANDARDIZE-VARIABLES(*rule*)
      for each θ such that SUBST($\theta, p_1 \wedge \ldots \wedge p_n$) = SUBST($\theta, p'_1 \wedge \ldots \wedge p'_n$)
          for some $p'_1, \ldots, p'_n$ in KB
        $q' \leftarrow$ SUBST(θ, q)
        if q' does not unify with some sentence already in KB or *new* **then**
          add q' to *new*
          $\phi \leftarrow$ UNIFY(q', α)
          if ϕ is not *failure* **then return** ϕ
    if $new = \{\}$ **then return** *false*
    add *new* to KB

Figure 9.3 A conceptually straightforward, but inefficient, forward-chaining algorithm. On each iteration, it adds to KB all the atomic sentences that can be inferred in one step from the implication sentences and the atomic sentences already in KB. The function STANDARDIZE-VARIABLES replaces all variables in its arguments with new ones that have not been used before.

function FOL-BC-ASK(KB, *query*) **returns** a generator of substitutions
  return FOL-BC-OR(KB, *query*, { })

function FOL-BC-OR(KB, *goal*, θ) **returns** a substitution
  for each *rule* in FETCH-RULES-FOR-GOAL(KB, *goal*) **do**
    (*lhs* $\Rightarrow$ *rhs*) $\leftarrow$ STANDARDIZE-VARIABLES(*rule*)
    for each θ' **in** FOL-BC-AND(KB, *lhs*, UNIFY(*rhs*, *goal*, θ)) **do**
      yield θ'

function FOL-BC-AND(KB, *goals*, θ) **returns** a substitution
  if θ = *failure* **then return**
  else if LENGTH(*goals*) = 0 **then yield** θ
  else
    first,*rest* $\leftarrow$ FIRST(*goals*), REST(*goals*)
    for each θ' **in** FOL-BC-OR(KB, SUBST(θ, *first*), θ) **do**
      for each θ'' **in** FOL-BC-AND(KB, *rest*, θ') **do**
        yield θ''

Figure 9.6 A simple backward-chaining algorithm for first-order knowledge bases.

procedure APPEND($ax, y, az, continuation$)

$trail \leftarrow$ GLOBAL-TRAIL-POINTER()
if ax = [ ] and UNIFY(y, az) **then** CALL($continuation$)
RESET-TRAIL($trail$)
$a, x, z \leftarrow$ NEW-VARIABLE(), NEW-VARIABLE(), NEW-VARIABLE()
if UNIFY($ax, [a] + x$) and UNIFY($az, [a \mid z]$) **then** APPEND($x, y, z, continuation$)

Figure 9.8 Pseudocode representing the result of compiling the `Append` predicate. The function NEW-VARIABLE returns a new variable, distinct from all other variables used so far. The procedure CALL($continuation$) continues execution with the specified continuation.

CHAPTER 10

KNOWLEDGE REPRESENTATION

CHAPTER 11

AUTOMATED PLANNING

$$
\begin{aligned}
& \mathit{Init}(\mathit{At}(C_1,\ \mathit{SFO}) \land \mathit{At}(C_2,\ \mathit{JFK}) \land \mathit{At}(P_1,\ \mathit{SFO}) \land \mathit{At}(P_2,\ \mathit{JFK}) \\
& \quad \land \mathit{Cargo}(C_1) \land \mathit{Cargo}(C_2) \land \mathit{Plane}(P_1) \land \mathit{Plane}(P_2) \\
& \quad \land \mathit{Airport}(\mathit{JFK}) \land \mathit{Airport}(\mathit{SFO})) \\
& \mathit{Goal}(\mathit{At}(C_1,\ \mathit{JFK}) \land \mathit{At}(C_2,\ \mathit{SFO})) \\
& \mathit{Action}(\mathit{Load}(c,\ p,\ a), \\
& \quad \text{PRECOND: } \mathit{At}(c,\ a) \land \mathit{At}(p,\ a) \land \mathit{Cargo}(c) \land \mathit{Plane}(p) \land \mathit{Airport}(a) \\
& \quad \text{EFFECT: } \neg\, \mathit{At}(c,\ a) \land \mathit{In}(c,\ p)) \\
& \mathit{Action}(\mathit{Unload}(c,\ p,\ a), \\
& \quad \text{PRECOND: } \mathit{In}(c,\ p) \land \mathit{At}(p,\ a) \land \mathit{Cargo}(c) \land \mathit{Plane}(p) \land \mathit{Airport}(a) \\
& \quad \text{EFFECT: } \mathit{At}(c,\ a) \land \neg\, \mathit{In}(c,\ p)) \\
& \mathit{Action}(\mathit{Fly}(p,\ \mathit{from},\ \mathit{to}), \\
& \quad \text{PRECOND: } \mathit{At}(p,\ \mathit{from}) \land \mathit{Plane}(p) \land \mathit{Airport}(\mathit{from}) \land \mathit{Airport}(\mathit{to}) \\
& \quad \text{EFFECT: } \neg\, \mathit{At}(p,\ \mathit{from}) \land \mathit{At}(p,\ \mathit{to}))
\end{aligned}
$$

Figure 11.1 A PDDL description of an air cargo transportation planning problem.

$$
\begin{aligned}
& \mathit{Init}(\mathit{Tire}(\mathit{Flat}) \land \mathit{Tire}(\mathit{Spare}) \land \mathit{At}(\mathit{Flat}, \mathit{Axle}) \land \mathit{At}(\mathit{Spare}, \mathit{Trunk})) \\
& \mathit{Goal}(\mathit{At}(\mathit{Spare}, \mathit{Axle})) \\
& \mathit{Action}(\mathit{Remove}(\mathit{obj}, \mathit{loc}), \\
& \quad \text{PRECOND: } \mathit{At}(\mathit{obj}, \mathit{loc}) \\
& \quad \text{EFFECT: } \neg\, \mathit{At}(\mathit{obj}, \mathit{loc}) \land \mathit{At}(\mathit{obj}, \mathit{Ground})) \\
& \mathit{Action}(\mathit{PutOn}(t,\ \mathit{Axle}), \\
& \quad \text{PRECOND: } \mathit{Tire}(t) \land \mathit{At}(t, \mathit{Ground}) \land \neg\, \mathit{At}(\mathit{Flat}, \mathit{Axle}) \land \neg\, \mathit{At}(\mathit{Spare}, \mathit{Axle}) \\
& \quad \text{EFFECT: } \neg\, \mathit{At}(t, \mathit{Ground}) \land \mathit{At}(t, \mathit{Axle})) \\
& \mathit{Action}(\mathit{LeaveOvernight}, \\
& \quad \text{PRECOND: } \\
& \quad \text{EFFECT: } \neg\, \mathit{At}(\mathit{Spare}, \mathit{Ground}) \land \neg\, \mathit{At}(\mathit{Spare}, \mathit{Axle}) \land \neg\, \mathit{At}(\mathit{Spare}, \mathit{Trunk}) \\
& \qquad\qquad \land \neg\, \mathit{At}(\mathit{Flat}, \mathit{Ground}) \land \neg\, \mathit{At}(\mathit{Flat}, \mathit{Axle}) \land \neg\, \mathit{At}(\mathit{Flat},\ \mathit{Trunk}))
\end{aligned}
$$

Figure 11.2 The simple spare tire problem.

$$\begin{array}{l}
\mathit{Init}(\mathit{On}(A,\mathit{Table}) \land \mathit{On}(B,\mathit{Table}) \land \mathit{On}(C,A) \\
\quad \land \mathit{Block}(A) \land \mathit{Block}(B) \land \mathit{Block}(C) \land \mathit{Clear}(B) \land \mathit{Clear}(C) \land \mathit{Clear}(\mathit{Table})) \\
\mathit{Goal}(\mathit{On}(A,B) \land \mathit{On}(B,C)) \\
\mathit{Action}(\mathit{Move}(b,x,y), \\
\quad \textsc{Precond}: \mathit{On}(b,x) \land \mathit{Clear}(b) \land \mathit{Clear}(y) \land \mathit{Block}(b) \land \mathit{Block}(y) \land \\
\qquad\qquad (b \neq x) \land (b \neq y) \land (x \neq y), \\
\quad \textsc{Effect}: \mathit{On}(b,y) \land \mathit{Clear}(x) \land \neg \mathit{On}(b,x) \land \neg \mathit{Clear}(y)) \\
\mathit{Action}(\mathit{MoveToTable}(b,x), \\
\quad \textsc{Precond}: \mathit{On}(b,x) \land \mathit{Clear}(b) \land \mathit{Block}(b) \land \mathit{Block}(x), \\
\quad \textsc{Effect}: \mathit{On}(b,\mathit{Table}) \land \mathit{Clear}(x) \land \neg \mathit{On}(b,x))
\end{array}$$

Figure 11.4 A planning problem in the blocks world: building a three-block tower. One solution is the sequence $[\mathit{MoveToTable}(C,A), \mathit{Move}(B,\mathit{Table},C), \mathit{Move}(A,\mathit{Table},B)]$.

$$\begin{array}{l}
\mathit{Refinement}(\mathit{Go}(\mathit{Home},\mathit{SFO}), \\
\quad \textsc{Steps}: [\mathit{Drive}(\mathit{Home},\mathit{SFOLongTermParking}), \\
\qquad\qquad \mathit{Shuttle}(\mathit{SFOLongTermParking},\mathit{SFO})]\ ) \\
\mathit{Refinement}(\mathit{Go}(\mathit{Home},\mathit{SFO}), \\
\quad \textsc{Steps}: [\mathit{Taxi}(\mathit{Home},\mathit{SFO})]\ ) \\
\\
\mathit{Refinement}(\mathit{Navigate}([a,b],[x,y]), \\
\quad \textsc{Precond}: a = x \land b = y \\
\quad \textsc{Steps}: [\,]\ ) \\
\mathit{Refinement}(\mathit{Navigate}([a,b],[x,y]), \\
\quad \textsc{Precond}: \mathit{Connected}([a,b],[a-1,b]) \\
\quad \textsc{Steps}: [\mathit{Left}, \mathit{Navigate}([a-1,b],[x,y])]\ ) \\
\mathit{Refinement}(\mathit{Navigate}([a,b],[x,y]), \\
\quad \textsc{Precond}: \mathit{Connected}([a,b],[a+1,b]) \\
\quad \textsc{Steps}: [\mathit{Right}, \mathit{Navigate}([a+1,b],[x,y])]\ ) \\
\ldots
\end{array}$$

Figure 11.7 Definitions of possible refinements for two high-level actions: going to San Francisco airport and navigating in the vacuum world. In the latter case, note the recursive nature of the refinements and the use of preconditions.

```
function HIERARCHICAL-SEARCH(problem, hierarchy) returns a solution or failure
  frontier ← a FIFO queue with [Act] as the only element
  while true do
      if IS-EMPTY(frontier) then return failure
      plan ← POP(frontier)          // chooses the shallowest plan in frontier
      hla ← the first HLA in plan, or null if none
      prefix,suffix ← the action subsequences before and after hla in plan
      outcome ← RESULT(problem.INITIAL, prefix)
      if hla is null then          // so plan is primitive and outcome is its result
          if problem.IS-GOAL(outcome) then return plan
      else for each sequence in REFINEMENTS(hla, outcome, hierarchy) do
          add APPEND(prefix, sequence, suffix) to frontier
```

Figure 11.8 A breadth-first implementation of hierarchical forward planning search. The initial plan supplied to the algorithm is [*Act*]. The REFINEMENTS function returns a set of action sequences, one for each refinement of the HLA whose preconditions are satisfied by the specified state, *outcome*.

```
function ANGELIC-SEARCH(problem, hierarchy, initialPlan) returns solution or fail
  frontier ← a FIFO queue with initialPlan as the only element
  while true do
    if EMPTY?(frontier) then return fail
    plan ← POP(frontier)          // chooses the shallowest node in frontier
    if REACH⁺(problem.INITIAL, plan) intersects problem.GOAL then
      if plan is primitive then return plan          // REACH⁺ is exact for primitive plans
      guaranteed ← REACH⁻(problem.INITIAL, plan) ∩ problem.GOAL
      if guaranteed≠{ } and MAKING-PROGRESS(plan, initialPlan) then
        finalState ← any element of guaranteed
        return DECOMPOSE(hierarchy, problem.INITIAL, plan, finalState)
      hla ← some HLA in plan
      prefix,suffix ← the action subsequences before and after hla in plan
      outcome ← RESULT(problem.INITIAL, prefix)
      for each sequence in REFINEMENTS(hla, outcome, hierarchy) do
        frontier ← Insert(APPEND(prefix, sequence, suffix), frontier)

function DECOMPOSE(hierarchy, s_0, plan, s_f) returns a solution
  solution ← an empty plan
  while plan is not empty do
    action ← REMOVE-LAST(plan)
    s_i ← a state in REACH⁻(s_0, plan) such that s_f ∈REACH⁻(s_i, action)
    problem ← a problem with INITIAL = s_i and GOAL = s_f
    solution ← APPEND(ANGELIC-SEARCH(problem, hierarchy, action), solution)
    s_f ← s_i
  return solution
```

Figure 11.11 A hierarchical planning algorithm that uses angelic semantics to identify and commit to high-level plans that work while avoiding high-level plans that don't. The predicate MAKING-PROGRESS checks to make sure that we aren't stuck in an infinite regression of refinements. At top level, call ANGELIC-SEARCH with $[Act]$ as the *initialPlan*.

$Jobs(\{AddEngine1 \prec AddWheels1 \prec Inspect1\},$
$\quad \{AddEngine2 \prec AddWheels2 \prec Inspect2\})$

$Resources(EngineHoists(1),\ WheelStations(1),\ Inspectors(e2),\ LugNuts(500))$

$Action(AddEngine1,\ \textsc{Duration}{:}30,$
$\quad \textsc{Use}{:}EngineHoists(1))$
$Action(AddEngine2,\ \textsc{Duration}{:}60,$
$\quad \textsc{Use}{:}EngineHoists(1))$
$Action(AddWheels1,\ \textsc{Duration}{:}30,$
$\quad \textsc{Consume}{:}LugNuts(20),\ \textsc{Use}{:}WheelStations(1))$
$Action(AddWheels2,\ \textsc{Duration}{:}15,$
$\quad \textsc{Consume}{:}LugNuts(20),\ \textsc{Use}{:}WheelStations(1))$
$Action(Inspect_i,\ \textsc{Duration}{:}10,$
$\quad \textsc{Use}{:}Inspectors(1))$

Figure 11.13 A job-shop scheduling problem for assembling two cars, with resource constraints. The notation $A \prec B$ means that action A must precede action B.

CHAPTER 12

QUANTIFYING UNCERTAINTY

function DT-AGENT($percept$) **returns** an $action$
  persistent: $belief\_state$, probabilistic beliefs about the current state of the world
    $action$, the agent's action

  update $belief\_state$ based on $action$ and $percept$
  calculate outcome probabilities for actions,
    given action descriptions and current $belief\_state$
  select $action$ with highest expected utility
    given probabilities of outcomes and utility information
  return $action$

Figure 12.1 A decision-theoretic agent that selects rational actions.

CHAPTER 13

PROBABILISTIC REASONING

function ENUMERATION-ASK(X, **e**, bn) **returns** a distribution over X
  inputs: X, the query variable
      e, observed values for variables **E**
      bn, a Bayes net with variables $vars$

  $\mathbf{Q}(X) \leftarrow$ a distribution over X, initially empty
  for each value x_i of X **do**
    $\mathbf{Q}(x_i) \leftarrow$ ENUMERATE-ALL($vars$, $\mathbf{e}_{x_i}$)
      where $\mathbf{e}_{x_i}$ is **e** extended with $X = x_i$
  return NORMALIZE($\mathbf{Q}(X)$)

function ENUMERATE-ALL($vars$, **e**) **returns** a real number
  if EMPTY?($vars$) **then return** 1.0
  $V \leftarrow$ FIRST($vars$)
  if V is an evidence variable with value v in **e**
    then return $P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), **e**)
    else return $\sum_v P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}_v$)
      where $\mathbf{e}_v$ is **e** extended with $V = v$

Figure 13.11 The enumeration algorithm for exact inference in Bayes nets.

function ELIMINATION-ASK(X, **e**, bn) **returns** a distribution over X
  inputs: X, the query variable
      e, observed values for variables **E**
      bn, a Bayesian network with variables $vars$

  $factors \leftarrow [\,]$
  for each V **in** ORDER($vars$) **do**
    $factors \leftarrow$ [MAKE-FACTOR(V, **e**)] $+$ $factors$
    if V is a hidden variable **then** $factors \leftarrow$ SUM-OUT(V, $factors$)
  return NORMALIZE(POINTWISE-PRODUCT($factors$))

Figure 13.13 The variable elimination algorithm for exact inference in Bayes nets.

function PRIOR-SAMPLE(bn) **returns** an event sampled from the prior specified by bn
  inputs: bn, a Bayesian network specifying joint distribution $\mathbf{P}(X_1, \ldots, X_n)$

  $\mathbf{x} \leftarrow$ an event with n elements
  for each variable X_i **in** $X_1, \ldots, X_n$ **do**
    $\mathbf{x}[i] \leftarrow$ a random sample from $\mathbf{P}(X_i \mid parents(X_i))$
  return $\mathbf{x}$

Figure 13.16 A sampling algorithm that generates events from a Bayesian network. Each variable is sampled according to the conditional distribution given the values already sampled for the variable's parents.

function REJECTION-SAMPLING(X, $\mathbf{e}$, bn, N) **returns** an estimate of $\mathbf{P}(X \mid \mathbf{e})$
  inputs: X, the query variable
    $\mathbf{e}$, observed values for variables $\mathbf{E}$
    bn, a Bayesian network
    N, the total number of samples to be generated
  local variables: $\mathbf{C}$, a vector of counts for each value of X, initially zero

  for j = 1 **to** N **do**
    $\mathbf{x} \leftarrow$ PRIOR-SAMPLE(bn)
    if $\mathbf{x}$ is consistent with $\mathbf{e}$ **then**
      $\mathbf{C}[j] \leftarrow \mathbf{C}[j]$+1 where x_j is the value of X in $\mathbf{x}$
  return NORMALIZE($\mathbf{C}$)

Figure 13.17 The rejection-sampling algorithm for answering queries given evidence in a Bayesian network.

function LIKELIHOOD-WEIGHTING(X, **e**, bn, N) **returns** an estimate of $\mathbf{P}(X \mid \mathbf{e})$
  inputs: X, the query variable
    e, observed values for variables **E**
    bn, a Bayesian network specifying joint distribution $\mathbf{P}(X_1, \ldots, X_n)$
    N, the total number of samples to be generated
  local variables: **W**, a vector of weighted counts for each value of X, initially zero

  for j = 1 **to** N **do**
    x, $w \leftarrow$ WEIGHTED-SAMPLE(bn, **e**)
    $\mathbf{W}[j] \leftarrow \mathbf{W}[j] + w$ where x_j is the value of X in **x**
  return NORMALIZE(**W**)

function WEIGHTED-SAMPLE(bn, **e**) **returns** an event and a weight

  $w \leftarrow 1$; $\mathbf{x} \leftarrow$ an event with n elements, with values fixed from **e**
  for i = 1 **to** n **do**
    if X_i is an evidence variable with value x_{ij} in **e**
      then $w \leftarrow w \times P(X_i = x_{ij} \mid parents(X_i))$
      else $\mathbf{x}[i] \leftarrow$ a random sample from $\mathbf{P}(X_i \mid parents(X_i))$
  return **x**, w

Figure 13.18 The likelihood-weighting algorithm for inference in Bayesian networks. In WEIGHTED-SAMPLE, each nonevidence variable is sampled according to the conditional distribution given the values already sampled for the variable's parents, while a weight is accumulated based on the likelihood for each evidence variable.

function GIBBS-ASK(X, **e**, bn, N) **returns** an estimate of $\mathbf{P}(X \mid \mathbf{e})$
  local variables: **C**, a vector of counts for each value of X, initially zero
    Z, the nonevidence variables in bn
    x, the current state of the network, initialized from **e**

  initialize **x** with random values for the variables in **Z**
  for k = 1 **to** N **do**
    choose any variable Z_i from **Z** according to any distribution $\rho(i)$
    set the value of Z_i in **x** by sampling from $\mathbf{P}(Z_i \mid mb(Z_i))$
    $\mathbf{C}[j] \leftarrow \mathbf{C}[j] + 1$ where x_j is the value of X in **x**
  return NORMALIZE(**C**)

Figure 13.20 The Gibbs sampling algorithm for approximate inference in Bayes nets; this version chooses variables at random, but cycling through the variables but also works.

CHAPTER 14

PROBABILISTIC REASONING OVER TIME

function FORWARD-BACKWARD(**ev**, $prior$) **returns** a vector of probability distributions
  inputs: **ev**, a vector of evidence values for steps $1, \ldots, t$
    $prior$, the prior distribution on the initial state, $\mathbf{P}(\mathbf{X}_0)$
  local variables: **fv**, a vector of forward messages for steps $0, \ldots, t$
    b, a representation of the backward message, initially all 1s
    sv, a vector of smoothed estimates for steps $1, \ldots, t$

  fv$[0] \leftarrow prior$
  for $i = 1$ **to** t **do**
    fv$[i] \leftarrow$ FORWARD(**fv**$[i-1]$, **ev**$[i]$)
  for $i = t$ **down to** 1 **do**
    sv$[i] \leftarrow$ NORMALIZE(**fv**$[i] \times$ **b**)
    b $\leftarrow$ BACKWARD(**b**, **ev**$[i]$)
  return sv

Figure 14.4 The forward–backward algorithm for smoothing: computing posterior probabilities of a sequence of states given a sequence of observations. The FORWARD and BACKWARD operators are defined by Equations (**??**) and (**??**), respectively.

function FIXED-LAG-SMOOTHING(e_t, hmm, d) **returns** a distribution over $\mathbf{X}_{t-d}$
  inputs: e_t, the current evidence for time step t
    hmm, a hidden Markov model with $S \times S$ transition matrix $\mathbf{T}$
    d, the length of the lag for smoothing
  persistent: t, the current time, initially 1
    $\mathbf{f}$, the forward message $\mathbf{P}(X_t \mid e_{1:t})$, initially hmm.PRIOR
    $\mathbf{B}$, the d-step backward transformation matrix, initially the identity matrix
    $e_{t-d:t}$, double-ended list of evidence from $t-d$ to t, initially empty
  local variables: $\mathbf{O}_{t-d}$, $\mathbf{O}_t$, diagonal matrices containing the sensor model information

  add e_t to the end of $e_{t-d:t}$
  $\mathbf{O}_t \leftarrow$ diagonal matrix containing $\mathbf{P}(e_t \mid X_t)$
  if $t > d$ **then**
    $\mathbf{f} \leftarrow$ FORWARD($\mathbf{f}$, e_{t-d})
    remove e_{t-d-1} from the beginning of $e_{t-d:t}$
    $\mathbf{O}_{t-d} \leftarrow$ diagonal matrix containing $\mathbf{P}(e_{t-d} \mid X_{t-d})$
    $\mathbf{B} \leftarrow \mathbf{O}_{t-d}^{-1}\mathbf{T}^{-1}\mathbf{B}\mathbf{T}\mathbf{O}_t$
  else $\mathbf{B} \leftarrow \mathbf{B}\mathbf{T}\mathbf{O}_t$
  $t \leftarrow t + 1$
  if $t > d + 1$ **then return** NORMALIZE($\mathbf{f} \times \mathbf{B1}$) **else return** null

Figure 14.6 An algorithm for smoothing with a fixed time lag of d steps, implemented as an online algorithm that outputs the new smoothed estimate given the observation for a new time step. Notice that the final output NORMALIZE($\mathbf{f} \times \mathbf{B1}$) is just $\alpha\,\mathbf{f} \times \mathbf{b}$, by Equation (**??**).

function PARTICLE-FILTERING($\mathbf{e}$, N, dbn) **returns** a set of samples for the next time step
  inputs: $\mathbf{e}$, the new incoming evidence
    N, the number of samples to be maintained
    dbn, a DBN defined by $\mathbf{P}(\mathbf{X}_0)$, $\mathbf{P}(\mathbf{X}_1 \mid \mathbf{X}_0)$, and $\mathbf{P}(\mathbf{E}_1 \mid \mathbf{X}_1)$
  persistent: S, a vector of samples of size N, initially generated from $\mathbf{P}(\mathbf{X}_0)$
  local variables: W, a vector of weights of size N

  for i = 1 to N **do**
    $S[i] \leftarrow$ sample from $\mathbf{P}(\mathbf{X}_1 \mid \mathbf{X}_0 = S[i])$ // *step 1*
    $W[i] \leftarrow \mathbf{P}(\mathbf{e} \mid \mathbf{X}_1 = S[i])$ // *step 2*
  $S \leftarrow$ WEIGHTED-SAMPLE-WITH-REPLACEMENT(N, S, W)   // *step 3*
  return S

Figure 14.17 The particle filtering algorithm implemented as a recursive update operation with state (the set of samples). Each of the sampling operations involves sampling the relevant slice variables in topological order, much as in PRIOR-SAMPLE. The WEIGHTED-SAMPLE-WITH-REPLACEMENT operation can be implemented to run in $O(N)$ expected time. The step numbers refer to the description in the text.

CHAPTER 15

PROBABILISTIC PROGRAMMING

type Researcher, Paper, Citation
random String Name(Researcher)
random String Title(Paper)
random Paper PubCited(Citation)
random String Text(Citation)
random Boolean Professor(Researcher)
origin Researcher Author(Paper)

$\#Researcher \sim OM(3,1)$
$Name(r) \sim NamePrior()$
$Professor(r) \sim Boolean(0.2)$
$\#Paper(Author = r) \sim$ **if** $Professor(r)$ **then** $OM(1.5, 0.5)$ **else** $OM(1, 0.5)$
$Title(p) \sim PaperTitlePrior()$
$CitedPaper(c) \sim UniformChoice(\{Paper\ p\})$
$Text(c) \sim HMMGrammar(Name(Author(CitedPaper(c))), Title(CitedPaper(c)))$

Figure 15.5 An OUPM for citation information extraction. For simplicity the model assumes one author per paper and omits details of the grammar and error models.

$$
\begin{array}{l}
\#\mathit{SeismicEvents} \sim \mathit{Poisson}(T * \lambda_e) \\
\mathit{Time}(e) \sim \mathit{UniformReal}(0, T) \\
\mathit{EarthQuake}(e) \sim \mathit{Boolean}(0.999) \\
\mathit{Location}(e) \sim \textbf{if } \mathit{Earthquake}(e) \textbf{ then } \mathit{SpatialPrior}() \textbf{ else } \mathit{UniformEarth}() \\
\mathit{Depth}(e) \sim \textbf{if } \mathit{Earthquake}(e) \textbf{ then } \mathit{UniformReal}(0, 700) \textbf{ else } \mathit{Exactly}(0) \\
\mathit{Magnitude}(e) \sim \mathit{Exponential}(\mathit{log}(10)) \\
\mathit{Detected}(e, p, s) \sim \mathit{Logistic}(\mathit{weights}(s, p), \mathit{Magnitude}(e),\ \mathit{Depth}(e),\ \mathit{Dist}(e, s)) \\
\#\mathit{Detections}(\mathit{site} = s) \sim \mathit{Poisson}(T * \lambda_f(s)) \\
\#\mathit{Detections}(\mathit{event}{=}e,\ \mathit{phase}{=}p,\ \mathit{station}{=}s) = \textbf{if } \mathit{Detected}(e, p, s) \textbf{ then } 1 \textbf{ else } 0 \\
\mathit{OnsetTime}(a, s) \textbf{ if } (\mathit{event}(a) = \mathit{null}) \textbf{ then } \sim \mathit{UniformReal}(0, T) \\
\quad \textbf{else} = \mathit{Time}(\mathit{event}(a)) + \mathit{GeoTT}(\mathit{Dist}(\mathit{event}(a), s), \mathit{Depth}(\mathit{event}(a)), \mathit{phase}(a)) \\
\qquad\qquad + \mathit{Laplace}(\mu_t(s), \sigma_t(s)) \\
\mathit{Amplitude}(a, s) \textbf{ if } (\mathit{event}(a) = \mathit{null}) \textbf{ then } \sim \mathit{NoiseAmpModel}(s) \\
\quad \textbf{else} = \mathit{AmpModel}(\mathit{Magnitude}(\mathit{event}(a)), \mathit{Dist}(\mathit{event}(a), s), \mathit{Depth}(\mathit{event}(a)), \mathit{phase}(a)) \\
\mathit{Azimuth}(a, s) \textbf{ if } (\mathit{event}(a) = \mathit{null}) \textbf{ then } \sim \mathit{UniformReal}(0,\ 360) \\
\quad \textbf{else} = \mathit{GeoAzimuth}(\mathit{Location}(\mathit{event}(a)), \mathit{Depth}(\mathit{event}(a)), \mathit{phase}(a), \mathit{Site}(s)) \\
\qquad\qquad + \mathit{Laplace}(0, \sigma_a(s)) \\
\mathit{Slowness}(a, s) \textbf{ if } (\mathit{event}(a) = \mathit{null}) \textbf{ then } \sim \mathit{UniformReal}(0, 20) \\
\quad \textbf{else} = \mathit{GeoSlowness}(\mathit{Location}(\mathit{event}(a)), \mathit{Depth}(\mathit{event}(a)), \mathit{phase}(a), \mathit{Site}(s)) \\
\qquad\qquad + \mathit{Laplace}(0, \sigma_s(s)) \\
\mathit{ObservedPhase}(a, s) \sim \mathit{CategoricalPhaseModel}(\mathit{phase}(a))
\end{array}
$$

Figure 15.6 A simplified version of the NET-VISA model (see text).

$$
\begin{array}{l}
\#\mathit{Aircraft}(\mathit{EntryTime} = t) \sim \mathit{Poisson}(\lambda_a) \\
\mathit{Exits}(a, t) \sim \textbf{if } \mathit{InFlight}(a, t) \textbf{ then } \mathit{Boolean}(\alpha_e) \\
\mathit{InFlight}(a, t) = (t{=}\mathit{EntryTime}(a)) \lor (\mathit{InFlight}(a, t-1) \land \neg\, \mathit{Exits}(a, t-1)) \\
X(a, t) \sim \textbf{if } t = \mathit{EntryTime}(a) \textbf{ then } \mathit{InitX}() \\
\quad \textbf{else if } \mathit{InFlight}(a, t) \textbf{ then } \mathcal{N}(\mathbf{F}\, X(a, t-1), \mathbf{\Sigma}_x) \\
\#\mathit{Blip}(\mathit{Source}{=}a,\ \mathit{Time}{=}t) \sim \textbf{if } \mathit{InFlight}(a, t) \textbf{ then } \mathit{Bernoulli}(\mathit{DetectionProb}(X(a, t))) \\
\#\mathit{Blip}(\mathit{Time}{=}t) \sim \mathit{Poisson}(\lambda_f) \\
Z(b) \sim \textbf{if } \mathit{Source}(b){=}\mathit{null} \textbf{ then } \mathit{UniformZ}(R) \textbf{ else } \mathcal{N}(\mathbf{H}\, X(\mathit{Source}(b), \mathit{Time}(b)), \mathbf{\Sigma}_z)
\end{array}
$$

Figure 15.9 An OUPM for radar tracking of multiple targets with false alarms, detection failure, and entry and exit of aircraft. The rate at which new aircraft enter the scene is λ_a, while the probability per time step that an aircraft exits the scene is α_e. False alarm blips (i.e., ones not produced by an aircraft) appear uniformly in space at a rate of λ_f per time step. The probability that an aircraft is detected (i.e., produces a blip) depends on its current position.

```
function GENERATE-IMAGE() returns an image with some letters
  letters ← GENERATE-LETTERS(10)
  return RENDER-NOISY-IMAGE(letters, 32, 128)

function GENERATE-LETTERS(λ) returns a vector of letters
  n ~ Poisson(λ)
  letters ← [ ]
  for i = 1 to n do
    letters[i] ~ UniformChoice({a, b, c, ···})
  return letters

function RENDER-NOISY-IMAGE(letters, width, height) returns a noisy image of the letters
  clean_image ← RENDER(letters, width, height, text_top = 10, text_left = 10)
  noisy_image ← [ ]
  noise_variance ~ UniformReal(0.1, 1)
  for row = 1 to width do
    for col = 1 to height do
      noisy_image[row, col] ~ N(clean_image[row, col], noise_variance)
  return noisy_image
```

Figure 15.11 Generative program for an open-universe probability model for optical character recognition. The generative program produces degraded images containing sequences of letters by generating each sequence, rendering it into a 2D image, and incorporating additive noise at each pixel.

```
function GENERATE-MARKOV-LETTERS(λ) returns a vector of letters
  n ~ Poisson(λ)
  letters ← [ ]
  letter_probs ← MARKOV-INITIAL()
  for i = 1 to n do
    letters[i] ~ Categorical(letter_probs)
    letter_probs ← MARKOV-TRANSITION(letters[i])
  return letters
```

Figure 15.15 Generative program for an improved optical character recognition model that generates letters according to a letter bigram model whose pairwise letter frequencies are estimated from a list of English words.

CHAPTER 16

MAKING SIMPLE DECISIONS

```
function INFORMATION-GATHERING-AGENT(percept) returns an action
   persistent: D, a decision network

   integrate percept into D
   j ← the value that maximizes VPI(E_j) / C(E_j)
   if VPI(E_j) > C(E_j)
      then return Request(E_j)
   else return the best action from D
```

Figure 16.9 Design of a simple, myopic information-gathering agent. The agent works by repeatedly selecting the observation with the highest information value, until the cost of the next observation is greater than its expected benefit.

CHAPTER 17

MAKING COMPLEX DECISIONS

function VALUE-ITERATION(mdp, ϵ) **returns** a utility function
  inputs: mdp, an MDP with states S, actions $A(s)$, transition model $P(s' \mid s, a)$,
      rewards $R(s, a, s')$, discount γ
    ϵ, the maximum error allowed in the utility of any state
  local variables: U, U', vectors of utilities for states in S, initially zero
      δ, the maximum relative change in the utility of any state

  repeat
    $U \leftarrow U'$; $\delta \leftarrow 0$
    for each state s **in** S **do**
      $U'[s] \leftarrow \max_{a \in A(s)}$ Q-VALUE(mdp, s, a, U)
      if $|U'[s] - U[s]| > \delta$ **then** $\delta \leftarrow |U'[s] - U[s]|$
  until $\delta \leq \epsilon(1-\gamma)/\gamma$
  return U

Figure 17.6 The value iteration algorithm for calculating utilities of states. The termination condition is from Equation (**??**).

function POLICY-ITERATION(mdp) **returns** a policy
  inputs: mdp, an MDP with states S, actions $A(s)$, transition model $P(s' \mid s, a)$
  local variables: U, a vector of utilities for states in S, initially zero
      π, a policy vector indexed by state, initially random

  repeat
    $U \leftarrow$ POLICY-EVALUATION(π, U, mdp)
    $unchanged?$ $\leftarrow$ true
    for each state s **in** S **do**
      $a^* \leftarrow \underset{a \in A(s)}{\text{argmax}}$ Q-VALUE(mdp, s, a, U)
      if Q-VALUE(mdp, s, a^*, U) $>$ Q-VALUE($mdp, s, \pi[s], U$) **then**
        $\pi[s] \leftarrow a^*$; $unchanged? \leftarrow$ false
  until $unchanged?$
  return π

Figure 17.9 The policy iteration algorithm for calculating an optimal policy.

function POMDP-VALUE-ITERATION($pomdp, \epsilon$) **returns** a utility function
  inputs: $pomdp$, a POMDP with states S, actions $A(s)$, transition model $P(s' \mid s, a)$,
      sensor model $P(e \mid s)$, rewards $R(s)$, discount γ
    ϵ, the maximum error allowed in the utility of any state
  local variables: U, U', sets of plans p with associated utility vectors α_p

  $U' \leftarrow$ a set containing just the empty plan [ ], with $\alpha_{[\,]}(s) = R(s)$
  repeat
    $U \leftarrow U'$
    $U' \leftarrow$ the set of all plans consisting of an action and, for each possible next percept,
      a plan in U with utility vectors computed according to Equation (**??**)
    $U' \leftarrow$ REMOVE-DOMINATED-PLANS(U')
  until MAX-DIFFERENCE(U, U') $\leq \epsilon(1-\gamma)/\gamma$
  return U

Figure 17.16 A high-level sketch of the value iteration algorithm for POMDPs. The REMOVE-DOMINATED-PLANS step and MAX-DIFFERENCE test are typically implemented as linear programs.

CHAPTER 18

MULTIAGENT DECISION MAKING

$$
\begin{array}{l}
Actors(A, B) \\
Init(At(A, LeftBaseline) \ \wedge \ At(B, RightNet) \ \wedge \\
\qquad Approaching(Ball, RightBaseline) \ \wedge \ Partner(A, B) \ \wedge \ Partner(B, A) \\
Goal(Returned(Ball) \ \wedge \ (At(x, RightNet) \ \vee \ At(x, LeftNet)) \\
Action(Hit(actor, Ball), \\
\qquad \text{PRECOND:} Approaching(Ball, loc) \ \wedge \ At(actor, loc) \\
\qquad \text{EFFECT:} Returned(Ball)) \\
Action(Go(actor, to), \\
\qquad \text{PRECOND:} At(actor, loc) \ \wedge \ to \neq loc, \\
\qquad \text{EFFECT:} At(actor, to) \ \wedge \ \neg \ At(actor, loc))
\end{array}
$$

Figure 18.1 The doubles tennis problem. Two actors, A and B, are playing together and can be in one of four locations: *LeftBaseline*, *RightBaseline*, *LeftNet*, and *RightNet*. The ball can be returned only if a player is in the right place. The *NoOp* action is a dummy, which has no effect. Note that each action must include the actor as an argument.

CHAPTER 19

LEARNING FROM EXAMPLES

function LEARN-DECISION-TREE($examples$, $attributes$, $parent\_examples$) **returns** a tree

  if $examples$ is empty **then return** PLURALITY-VALUE($parent\_examples$)
  else if all $examples$ have the same classification **then return** the classification
  else if $attributes$ is empty **then return** PLURALITY-VALUE($examples$)
  else
    $A \leftarrow \text{argmax}_{a \in attributes}$ IMPORTANCE(a, $examples$)
    $tree \leftarrow$ a new decision tree with root test A
    for each value v of A **do**
      $exs \leftarrow \{e : e \in examples$ **and** $e.A = v\}$
      $subtree \leftarrow$ LEARN-DECISION-TREE(exs, $attributes - A$, $examples$)
      add a branch to $tree$ with label $(A = v)$ and subtree $subtree$
    return $tree$

Figure 19.5 The decision tree learning algorithm. The function IMPORTANCE is described in Section **??**. The function PLURALITY-VALUE selects the most common output value among a set of examples, breaking ties randomly.

function MODEL-SELECTION($Learner$, $examples$, k) **returns** a (hypothesis, error rate) pair

```
err ← an array, indexed by size, storing validation-set error rates
training_set, test_set ← a partition of examples into two sets
for size = 1 to ∞ do
    err[size] ← CROSS-VALIDATION(Learner, size, training_set, k)
    if err is starting to increase significantly then
        best_size ← the value of size with minimum err[size]
        h ← Learner(best_size, training_set)
        return h, ERROR-RATE(h, test_set)
```

function CROSS-VALIDATION($Learner$, $size$, $examples$, k) **returns** error rate

```
N ← the number of examples
errs ← 0
for i = 1 to k do
    validation_set ← examples[(i − 1) × N/k:i × N/k]
    training_set ← examples − validation_set
    h ← Learner(size, training_set)
    errs ← errs + ERROR-RATE(h, validation_set)
return errs / k          // average error rate on validation sets, across k-fold cross-validation
```

Figure 19.8 An algorithm to select the model that has the lowest validation error. It builds models of increasing complexity, and choosing the one with best empirical error rate, err, on the validation data set. $Learner(size, examples)$ returns a hypothesis whose complexity is set by the parameter $size$, and which is trained on $examples$. In CROSS-VALIDATION, each iteration of the **for** loop selects a different slice of the $examples$ as the validation set, and keeps the other examples as the training set. It then returns the average validation set error over all the folds. Once we have determined which value of the $size$ parameter is best, MODEL-SELECTION returns the model (i.e., learner/hypothesis) of that size, trained on all the training examples, along with its error rate on the held-out test examples.

function DECISION-LIST-LEARNING($examples$) **returns** a decision list, or $failure$

```
if examples is empty then return the trivial decision list No
t ← a test that matches a nonempty subset examples_t of examples
     such that the members of examples_t are all positive or all negative
if there is no such t then return failure
if the examples in examples_t are positive then o ← Yes else o ← No
return a decision list with initial test t and outcome o and remaining tests given by
       DECISION-LIST-LEARNING(examples − examples_t)
```

Figure 19.11 An algorithm for learning decision lists.

function ADABOOST($examples, L, K$) **returns** a hypothesis
  inputs: $examples$, set of N labeled examples $(x_1, y_1), \ldots, (x_N, y_N)$
    L, a learning algorithm
    K, the number of hypotheses in the ensemble
  local variables: **w**, a vector of N example weights, initially all $1/N$
    h, a vector of K hypotheses
    z, a vector of K hypothesis weights

  $\epsilon \leftarrow$ a small positive number, used to avoid division by zero
  for k = 1 **to** K **do**
    $\mathbf{h}[k] \leftarrow L(examples, \mathbf{w})$
    $error \leftarrow 0$
    for j = 1 **to** N **do**   // *Compute the total error for* $\mathbf{h}[k]$
      if $\mathbf{h}[k](x_j) \neq y_j$ **then** $error \leftarrow error + \mathbf{w}[j]$
    if $error > 1/2$ **then break** from loop
    $error \leftarrow \min(error,\ 1 - \epsilon)$
    for j = 1 **to** N **do**   // *Give more weight to the examples* $\mathbf{h}[k]$ *got wrong*
      if $\mathbf{h}[k](x_j) = y_j$ **then** $\mathbf{w}[j] \leftarrow \mathbf{w}[j] \cdot\ error/(1 - error)$
    $\mathbf{w} \leftarrow$ NORMALIZE($\mathbf{w}$)
    $\mathbf{z}[k] \leftarrow \frac{1}{2}\ \log\left((1 - error)/error\right)$ // *Give more weight to accurate* $\mathbf{h}[k]$
  return $Function(x) :\ \sum \mathbf{z}_i\ \mathbf{h}_i(x)$

Figure 19.25 The ADABOOST variant of the boosting method for ensemble learning. The algorithm generates hypotheses by successively reweighting the training examples. The function WEIGHTED-MAJORITY generates a hypothesis that returns the output value with the highest vote from the hypotheses in **h**, with votes weighted by **z**. For regression problems, or for binary classification with two classes -1 and 1, this is $\sum_k \mathbf{h}[k]\mathbf{z}[k]$.

CHAPTER 20

LEARNING PROBABILISTIC MODELS

CHAPTER 21

DEEP LEARNING

CHAPTER 22

REINFORCEMENT LEARNING

```
function PASSIVE-ADP-LEARNER(percept) returns an action
  inputs: percept, a percept indicating the current state s′ and reward signal r
  persistent: π, a fixed policy
              mdp, an MDP with model P, rewards R, actions A, discount γ
              U, a table of utilities for states, initially empty
              N_{s′|s,a}, a table of outcome count vectors indexed by state and action, initially zero
              s, a, the previous state and action, initially null

  if s′ is new then U[s′] ← 0
  if s is not null then
    increment N_{s′|s,a}[s, a][s’]
    R[s, a, s′] ← r
    add a to A[s]
    P(· | s, a) ← NORMALIZE(N_{s′|s,a}[s, a])
    U ← POLICYEVALUATION(π, U, mdp)
    s, a ← s′, π[s′]
    return a
```

Figure 22.2 A passive reinforcement learning agent based on adaptive dynamic programming. The agent chooses a value for γ and then incrementally computes the P and R values of the MDP. The POLICY-EVALUATION function solves the fixed-policy Bellman equations, as described on page **??**.

function PASSIVE-TD-LEARNER(*percept*) **returns** an action
  inputs: *percept*, a percept indicating the current state s' and reward signal r
  persistent: π, a fixed policy
    s, the previous state, initially null
    U, a table of utilities for states, initially empty
    N_s, a table of frequencies for states, initially zero

  if s' is new **then** $U[s'] \leftarrow 0$
  if s is not null **then**
    increment $N_s[s]$
    $U[s] \leftarrow U[s] + \alpha(N_s[s]) \times (r + \gamma\, U[s']$ - $U[s])$
  $s \leftarrow s'$
  return $\pi[s']$

Figure 22.4 A passive reinforcement learning agent that learns utility estimates using temporal differences. The step-size function $\alpha(n)$ is chosen to ensure convergence.

function Q-LEARNING-AGENT(*percept*) **returns** an action
  inputs: *percept*, a percept indicating the current state s' and reward signal r
  persistent: Q, a table of action values indexed by state and action, initially zero
    N_{sa}, a table of frequencies for state–action pairs, initially zero
    s, a, the previous state and action, initially null

  if s is not null **then**
    increment $N_{sa}[s, a]$
    $Q[s, a] \leftarrow Q[s, a] + \alpha(N_{sa}[s, a])(r + \gamma \max_{a'} Q[s', a'] - Q[s, a])$
  $s, a \leftarrow s', \operatorname{argmax}_{a'} f(Q[s', a'], N_{sa}[s', a'])$
  return a

Figure 22.8 An exploratory Q-learning agent. It is an active learner that learns the value $Q(s, a)$ of each action in each situation. It uses the same exploration function f as the exploratory ADP agent, but avoids having to learn the transition model.

CHAPTER 23

NATURAL LANGUAGE PROCESSING

```
function CYK-PARSE(words, grammar) returns a table of parse trees
  inputs: words, a list of words
          grammar, a structure with LEXICALRULES and GRAMMARRULES
  T ← a table        // T[X, i, k] is most probable X tree spanning words_{i:k}
  P ← a table, initially all 0        // P[X, i, k] is probability of tree T[X, i, k]
  // Insert lexical categories for each word.
  for i = 1 to LEN(words) do
    for each (X, p) in grammar.LEXICALRULES(words_i) do
      P[X, i, i] ← p
      T[X, i, i] ← TREE(X, words_i)
  // Construct X_{i:k} from Y_{i:j} + Z_{j+1:k}, shortest spans first.
  for each (i, j, k) in SUBSPANS(LEN(words)) do
    for each (X, Y, Z, p) in grammar.GRAMMARRULES do
      PYZ ← P[Y, i, j] × P[Z, j + 1, k] × p
      if PYZ > P[X, i, k] do
        P[X, i, k] ← PYZ
        T[X, i, k] ← TREE(X, T[Y, i, j], T[Z, j + 1, k])
  return T

function SUBSPANS(N) yields (i, j, k) tuples
  for length = 2 to N do
    for i = 1 to N + 1 − length do
      k ← i + length − 1
      for j = i to k − 1 do
        yield (i, j, k)
```

Figure 23.5 The CYK algorithm for parsing. Given a sequence of words, it finds the most probable parse tree for the sequence and its subsequences. The table $P[X, i, k]$ gives the probability of the most probable tree of category X spanning $words_{i:k}$. The output table $T[X,\ i,\ k]$ contains the most probable tree of category X spanning positions i to k inclusive. The function SUBSPANS returns all tuples (i, j, k) covering a span of $words_{i:k}$, with $i \leq j < k$, listing the tuples by increasing length of the $i : k$ span, so that when we go to combine two shorter spans into a longer one, the shorter spans are already in the table. LEXICALRULES($word$) returns a collection of $(X,\ p)$ pairs, one for each rule of the form $X \rightarrow word$ [htbp], and GRAMMARRULES gives (X, Y, Z, p) tuples, one for each grammar rule of the form $X \rightarrow Y\ Z\ [p]$.

[[S [NP-2 **Her eyes**]
  [VP **were**
    [VP **glazed**
      [NP *-2]
      [$SBAR$-ADV **as if**
        [S [NP **she**]
          [VP **did n't**
            [VP [VP **hear** [NP *-1]]
              or
              [VP [$ADVP$ **even**] **see** [NP *-1]]
              [NP-1 **him**]]]]]]]]
 .]

Figure 23.8 Annotated tree for the sentence "Her eyes were glazed as if she didn't hear or even see him." from the Penn Treebank. Note a grammatical phenomenon we have not covered yet: the movement of a phrase from one part of the tree to another. This tree analyzes the phrase "hear or even see him" as consisting of two constituent VPs, [VP **hear** [NP *-1]] and [VP [$ADVP$ **even**] **see** [NP *-1]], both of which have a missing object, denoted *-1, which refers to the NP labeled elsewhere in the tree as [NP-1 **him**]. Similarly, the [NP *-2] refers to the [NP-2 **Her eyes**].

CHAPTER 24

DEEP LEARNING FOR NATURAL LANGUAGE PROCESSING

It is a truth universally acknowledged that the earth is not the center of the universe. There are those who assert there is. I do not accept them, but others I consider to be of the same opinion. The truth is, however, that if there are other than the center, and if there are any other living things in the universe and if they are not human, then we do not yet have our answers. We have to go on. This page gives a simplified, simplified answer to the problem. We don't have all the answers. The truth is, however, that the truth is out there.

When Gregor Samsa woke up one morning, he did not notice anything strange. "When my wife is looking at me, I feel like she is looking at a piece of art," he said. "I think she is admiring something I have created." The idea is that by looking at your own life, you learn something important and become a better person. It is a theory that emerged from psychologist Daniel Goleman's work, in which he asked "How do you know you're not a loser?"

Alice was beginning to get very tired of sitting with her sister on the bank. She sat up, yawned, and said, with a loud little scream, "I hope you don't mind if I keep on doing what I should like to do, and if someone asks me which of us will do more, don't tell them that I won't do much, my dear sister."

All happy families are alike; each happy family is like a garden of paradise. The only difference between happy families and unhappy families, is that the unhappy family doesn't have any flowers or trees.

Tell me a story. Tell me a story. Tell me a story. Tell me a story. Tell me a story. Tell me a story. Tell me a story. Tell me a story. Tell me a story. Tell me a story. Tell me a story. Tell me a story. Please fill out the following details. Thank you... Thank you for your interest in this interview. Please wait...

Figure 24.13 Example completion texts generated by the GPT-2 language model, given the prompts in **bold**. Most of the texts are quite fluent English, at least locally. The final example demonstrates that sometimes the model just breaks down.

CHAPTER 25

COMPUTER VISION

CHAPTER 26

ROBOTICS

function MONTE-CARLO-LOCALIZATIONa, z, N, $P(X'|X,\ v,\ \omega)$, $P(z|z^*)$, map
  returns a set of samples, S, for the next time step
  inputs: a, robot velocities v and ω
      z, a vector of M range scan data points
      $P(X'|X,\ v,\ \omega)$, motion model
      $P(z|z^*)$, a range sensor noise model
      map, a 2D map of the environment
  persistent: S, a vector of N samples
  local variables: W, a vector of N weights
      S', a temporary vector of N samples

  if S is empty **then**
    for $i = 1$ to N **do**   *// initialization phase*
      $S[i] \leftarrow$ sample from $P(X_0)$
  for $i = 1$ to N **do**   *// update cycle*
    $S'[i] \leftarrow$ sample from $P(X'|X = S[i], v, \omega)$
    $W[i] \leftarrow 1$
    for $j = 1$ to M **do**
      $z^* \leftarrow$ RAYCAST(j, $X = S'[i]$, map)
      $W[i] \leftarrow W[i] \cdot P(z_j|\, z^*)$
    $S \leftarrow$ WEIGHTED-SAMPLE-WITH-REPLACEMENT(N, S', W)
  return S

Figure 26.6 A Monte Carlo localization algorithm using a range-scan sensor model with independent noise.

CHAPTER 27

PHILOSOPHY, ETHICS, AND SAFETY OF AI

CHAPTER 28

THE FUTURE OF AI

CHAPTER 29

MATHEMATICAL BACKGROUND

CHAPTER 30

NOTES ON LANGUAGES AND ALGORITHMS