CSC242: Introduction to Artificial Intelligence

Lecture 2.2

Please put away all electronic devices

Constraint Satisfaction Problem (CSP)

- X: Set of variables $\{ X_1, ..., X_n \}$
- D: Set of domains $\{ D_1, ..., D_n \}$
  - Each D_i : set of values $\{ v_1, ..., v_k \}$
- C: Set of constraints $\{ C_1, ..., C_m \}$
- Solution: Assign to each X_i a value from D_i such that all the C_j are satisfied

Factored Representation

- Splits a state into factors (attributes, features, variables) that can have values
- Factored states can be more or less similar (unlike atomic states)
- Can also represent uncertainty (don't know the value of some attribute)

Backtracking Search for CSPs

- DFS search through the space of assignments
- Assign one variable at a time
- Because the representation of CSPs is standardized, no need to supply initial state, actions, transition model, or goal test!
- Early pruning of inconsistent states

Constraint Propagation

- Using the constraints to reduce the set of legal values of a variable, which can in turn reduce the legal values of another variable, and so on
- Not a search process!
- Part of state update in state-space search
- A type of <u>inference</u>: making implicit information explicit

Constraint Propagation

- Node consistency:
  - Propagate unary constraints (once)
- Arc consistency
  - Propagate binary constraints
  - AC-3 algorithm

Constraint Propagation

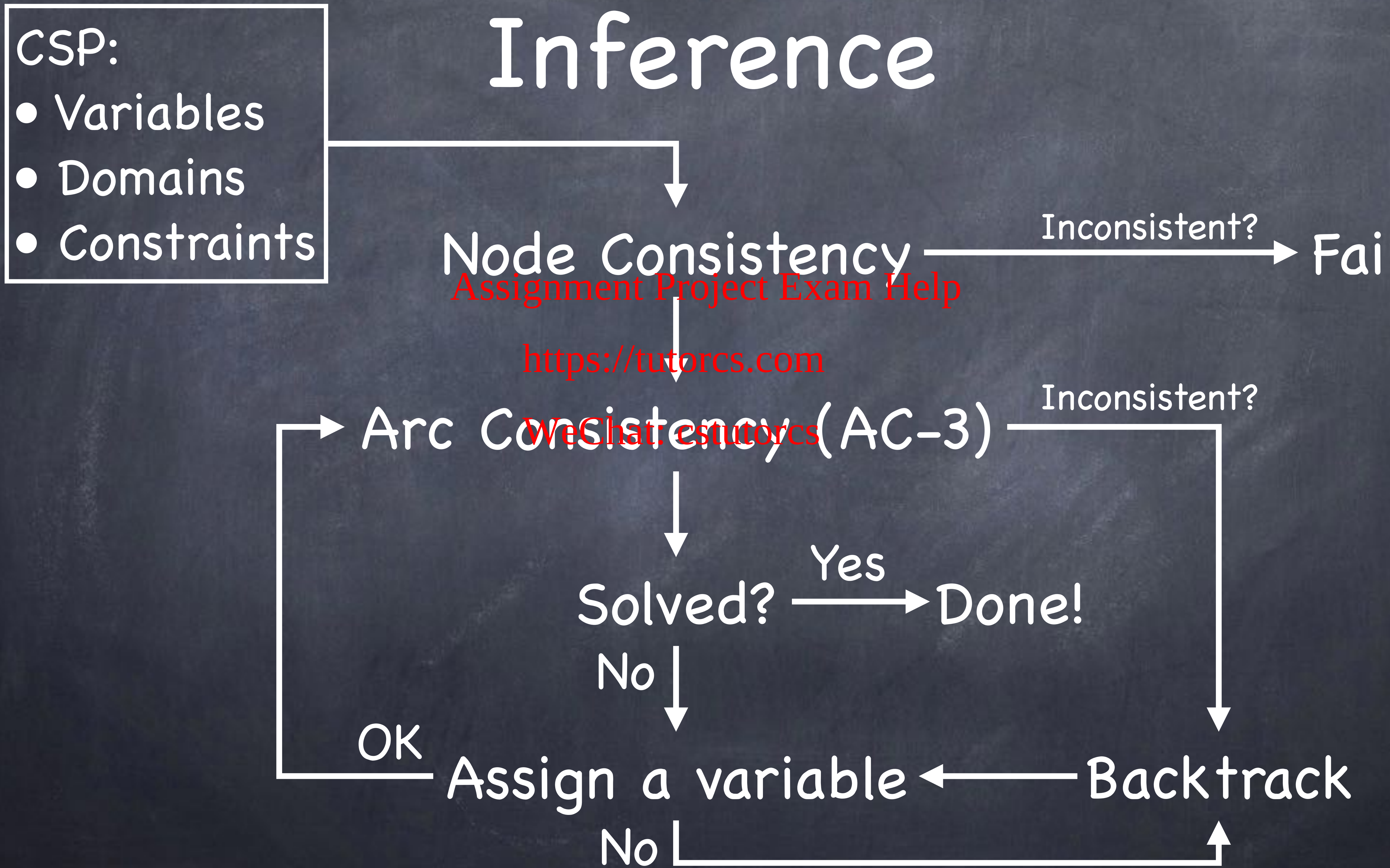
- "After constraint propagation, we are left with a CSP that is equivalent to the original CSP—they both have the same solutions—but the new CSP will in most cases be faster to search because its variables have smaller domains."

Constraint Propagation (inference) ⟷ State-Space Search

Interleaving Search and Inference

- After each choice during search, we can perform inference to reduce future search

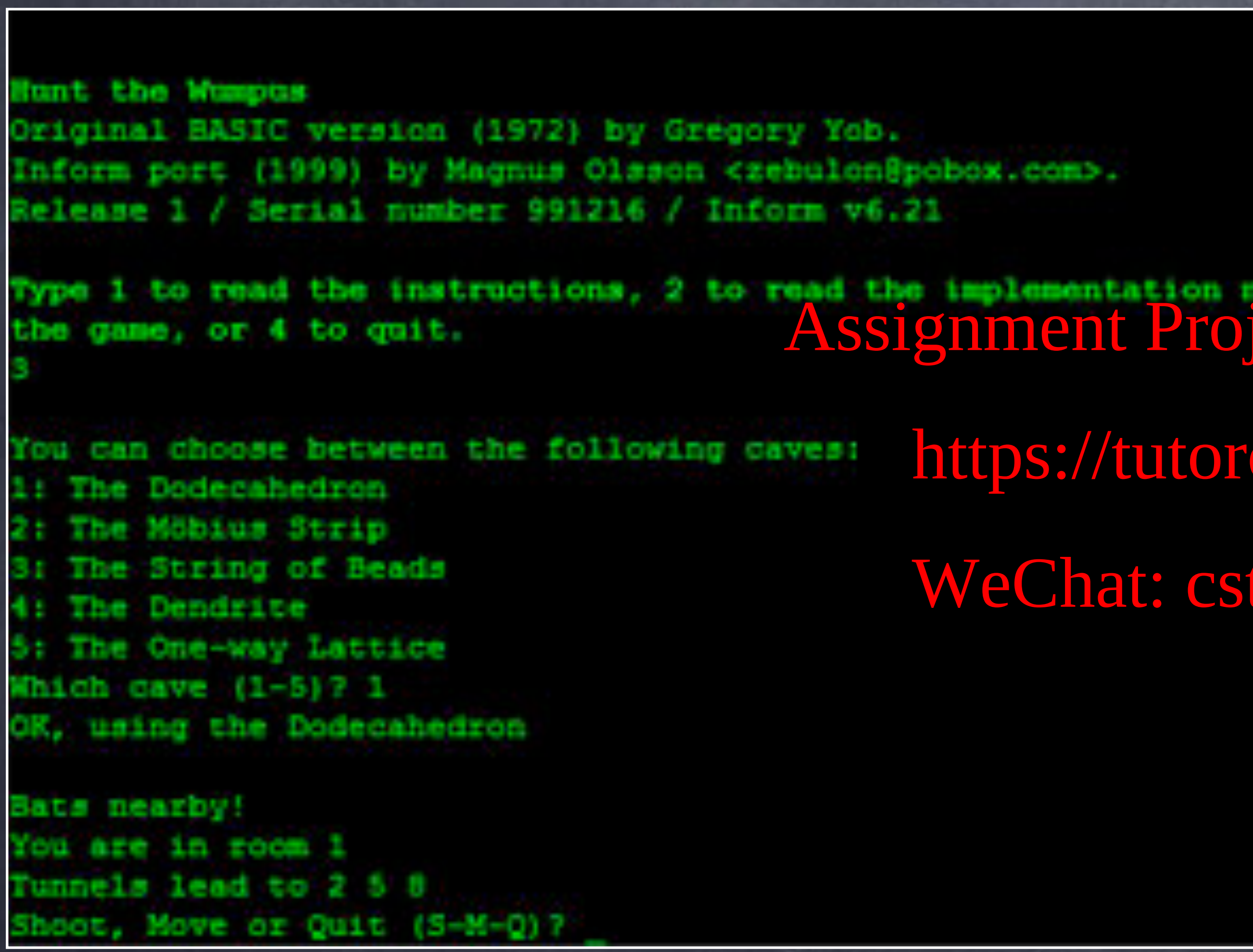
Interleaving Search and Inference

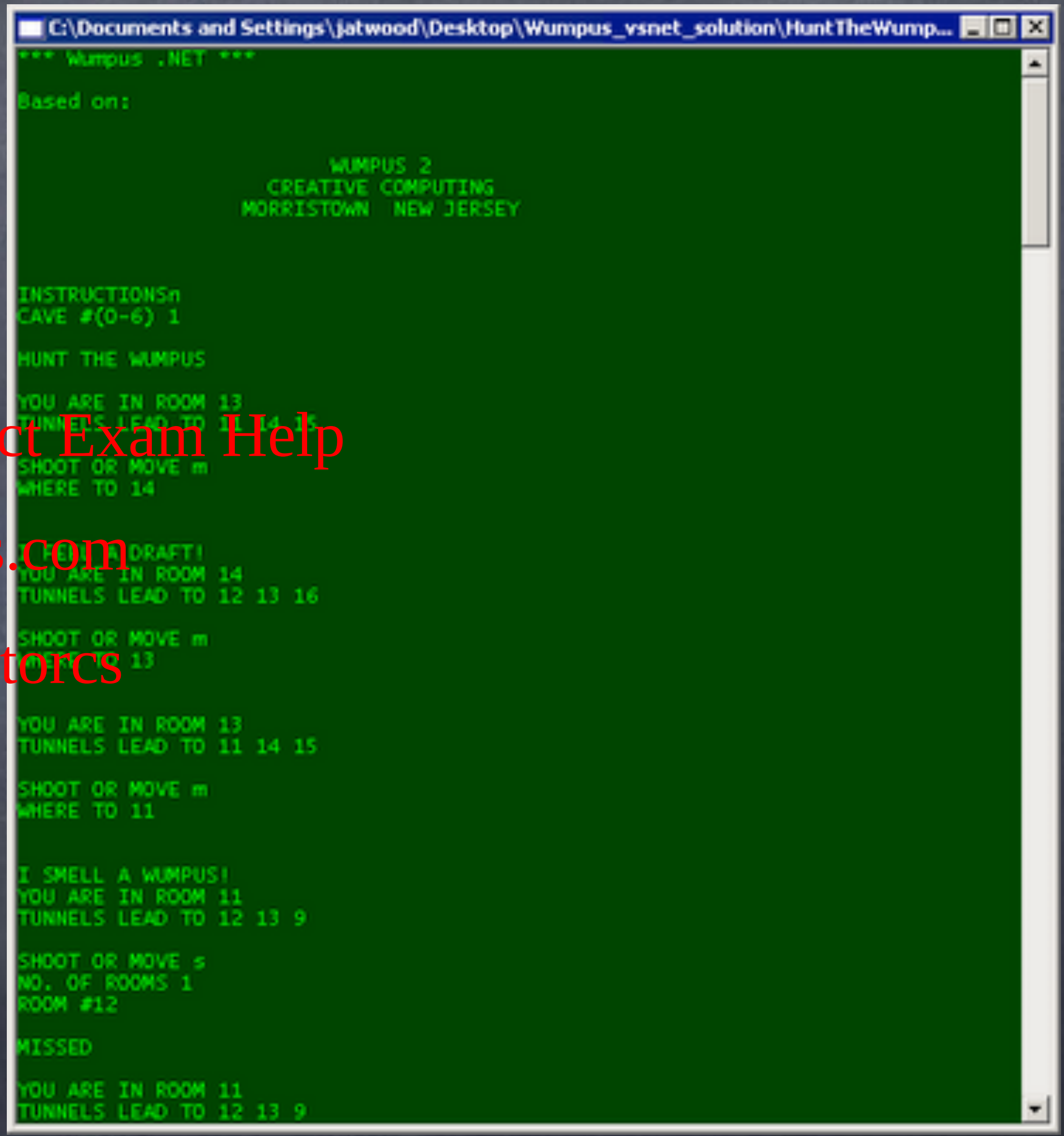


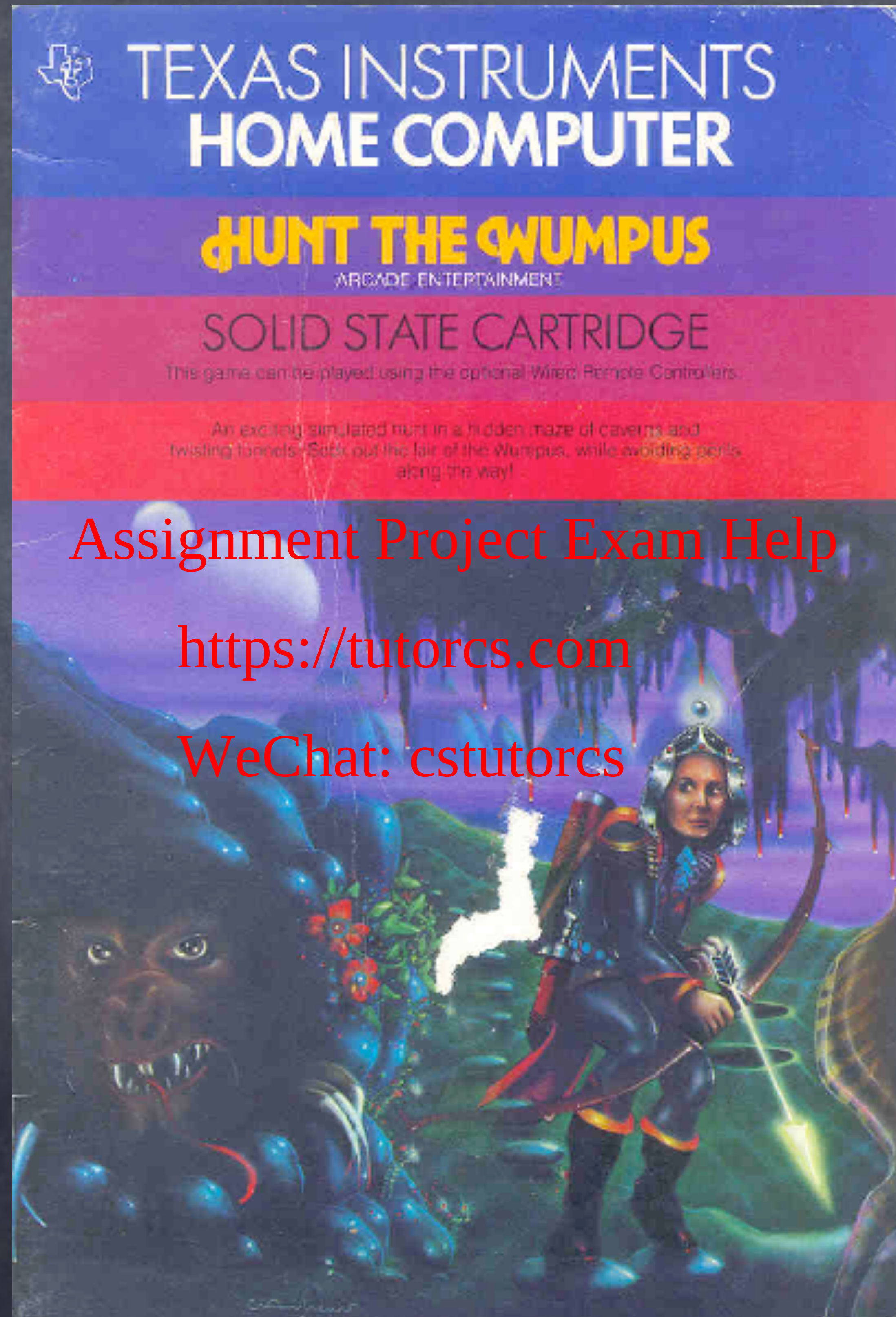
Constraint Satisfaction

- Impose a structure on the representation of states: Variables, Domains, Constraints
- Backtracking (DFS) search for complete, consistent assignment of values to variables
- Inference (constraint propagation) can reduce the domains of variables
  - Preprocessing and/or interleaved with search
- Useful problem-independent heuristics

Hunt The Wumpus


Hunt the Wumpus
Original BASIC version (1972) by Gregory Yob.
Inform port (1999) by Magnus Olsson <zebulon@pobox.com>.
Release 1 / Serial number 991216 / Inform v6.21
Type 1 to read the instructions, 2 to read the implementation n
the game, or 4 to quit.
3
You can choose between the following caves:
1: The Dodecahedron
2: The Möbius Strip
3: The String of Beads
4: The Dendrite
5: The One-way Lattice
Which cave (1-5)? 1
OK, using the Dodecahedron
Bats nearby!
You are in room 1
Tunnels lead to 2 5 8
Shoot, Move or Quit (S-M-Q)?

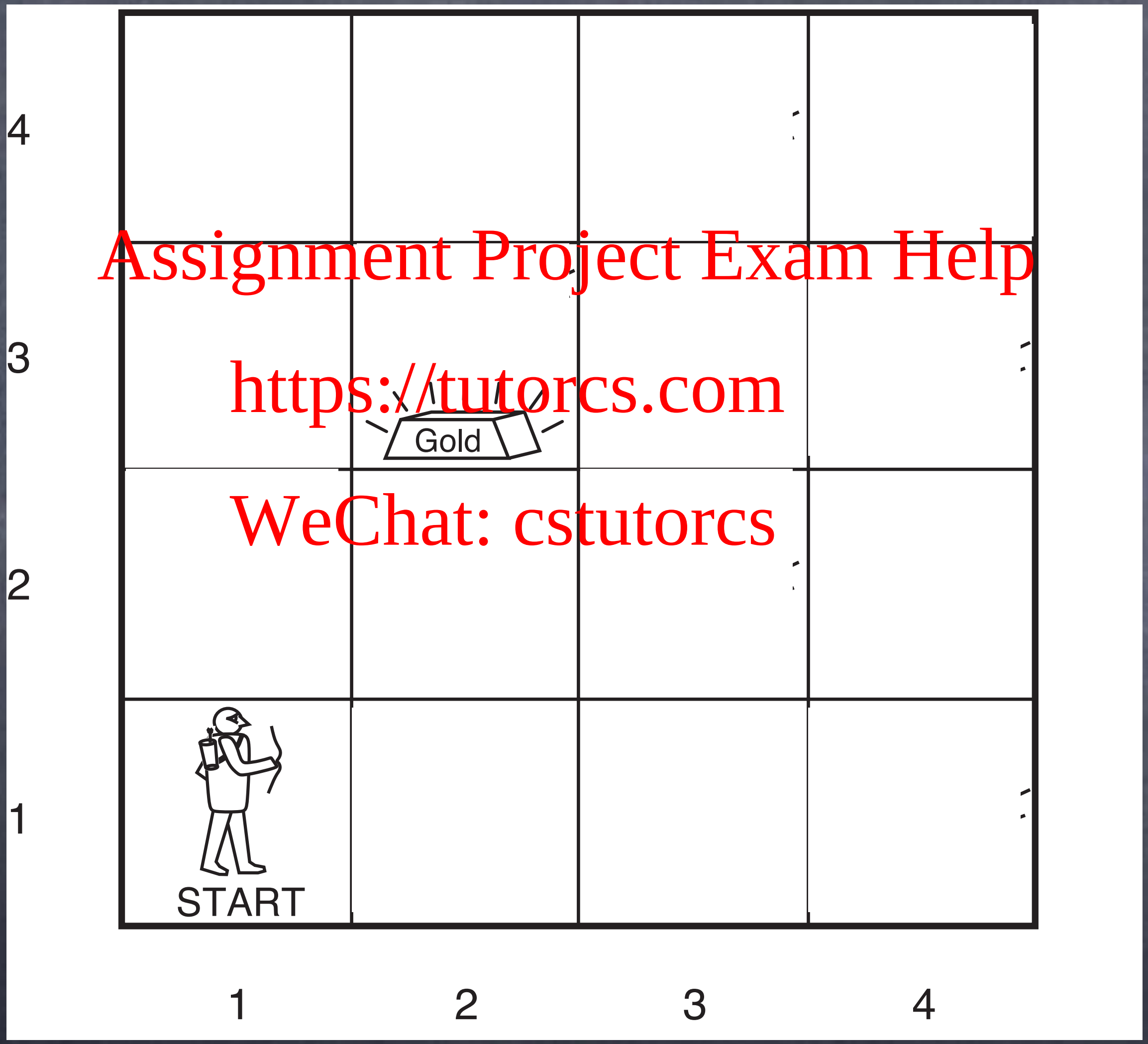

C:\Documents and Settings\jatwood\Desktop\Wumpus_vsnet_solution\HuntTheWump...
*** Wumpus .NET ***
Based on:
WUMPUS 2
CREATIVE COMPUTING
MORRISTOWN NEW JERSEY
INSTRUCTIONSn
CAVE #(0-6) 1
HUNT THE WUMPUS
YOU ARE IN ROOM 13
SHOOT OR MOVE m
WHERE TO 14
YOU ARE IN ROOM 14
TUNNELS LEAD TO 12 13 16
SHOOT OR MOVE m
YOU ARE IN ROOM 13
TUNNELS LEAD TO 11 14 15
SHOOT OR MOVE m
WHERE TO 11
I SMELL A WUMPUS!
YOU ARE IN ROOM 11
TUNNELS LEAD TO 12 13 9
SHOOT OR MOVE s
NO. OF ROOMS 1
ROOM #12
MISSED
YOU ARE IN ROOM 11
TUNNELS LEAD TO 12 13 9


TEXAS INSTRUMENTS
HOME COMPUTER
HUNT THE WUMPUS
ARCADE ENTERTAINMENT
SOLID STATE CARTRIDGE
This game can be played using the optional Wired Remote Controllers.
An exciting simulated hunt in a hidden maze of caverns and twisting tunnels. Seek out the lair of the Wumpus, while avoiding perils along the way!

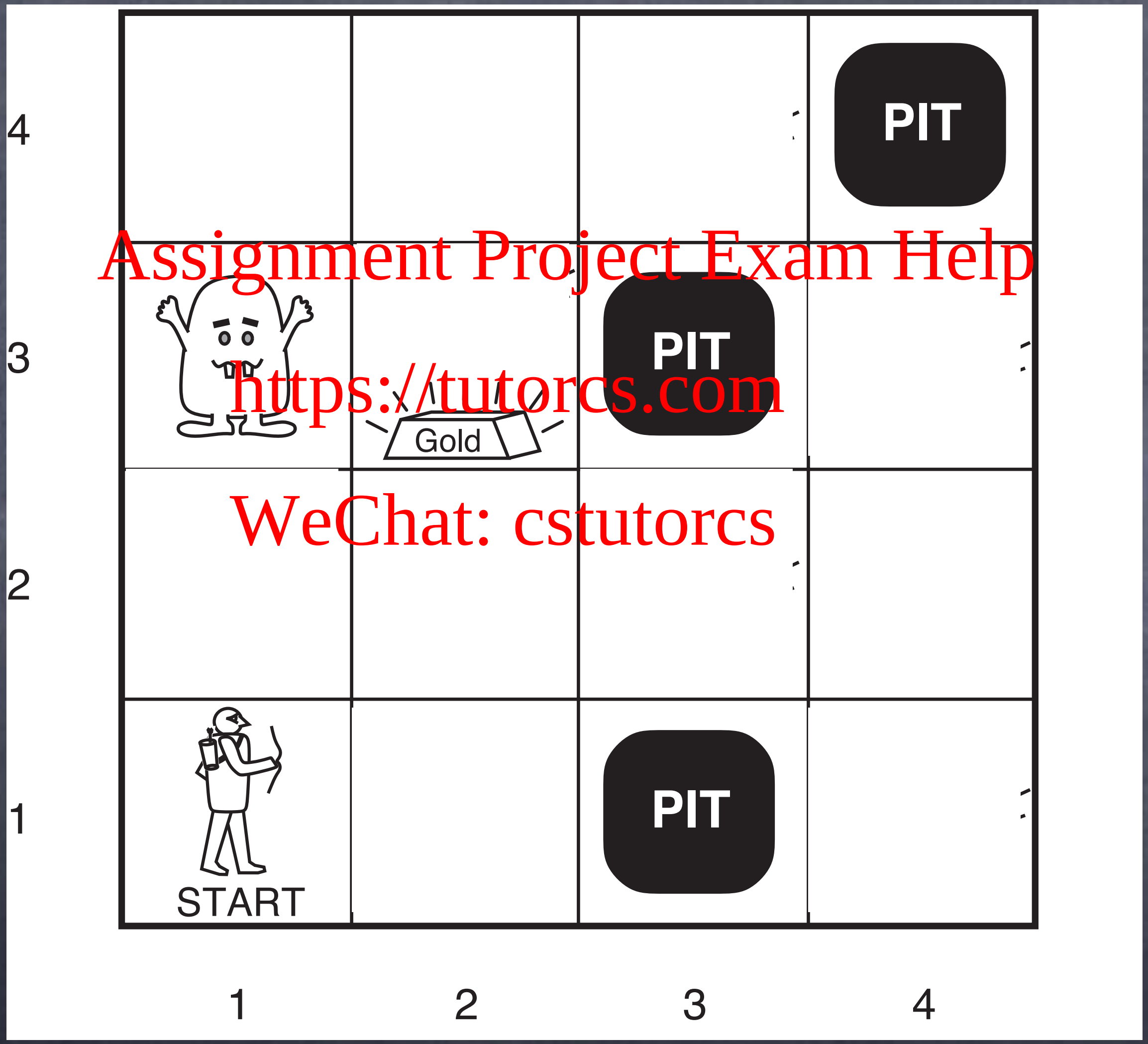

Hunt the Wumpus


Hunted Wumpus
3
Creature — Beast
When Hunted Wumpus comes into play, each other player may put a creature card from his or her hand into play.
Just one can feed a dozen people for a month.
Thomas M. Baxa
6/6

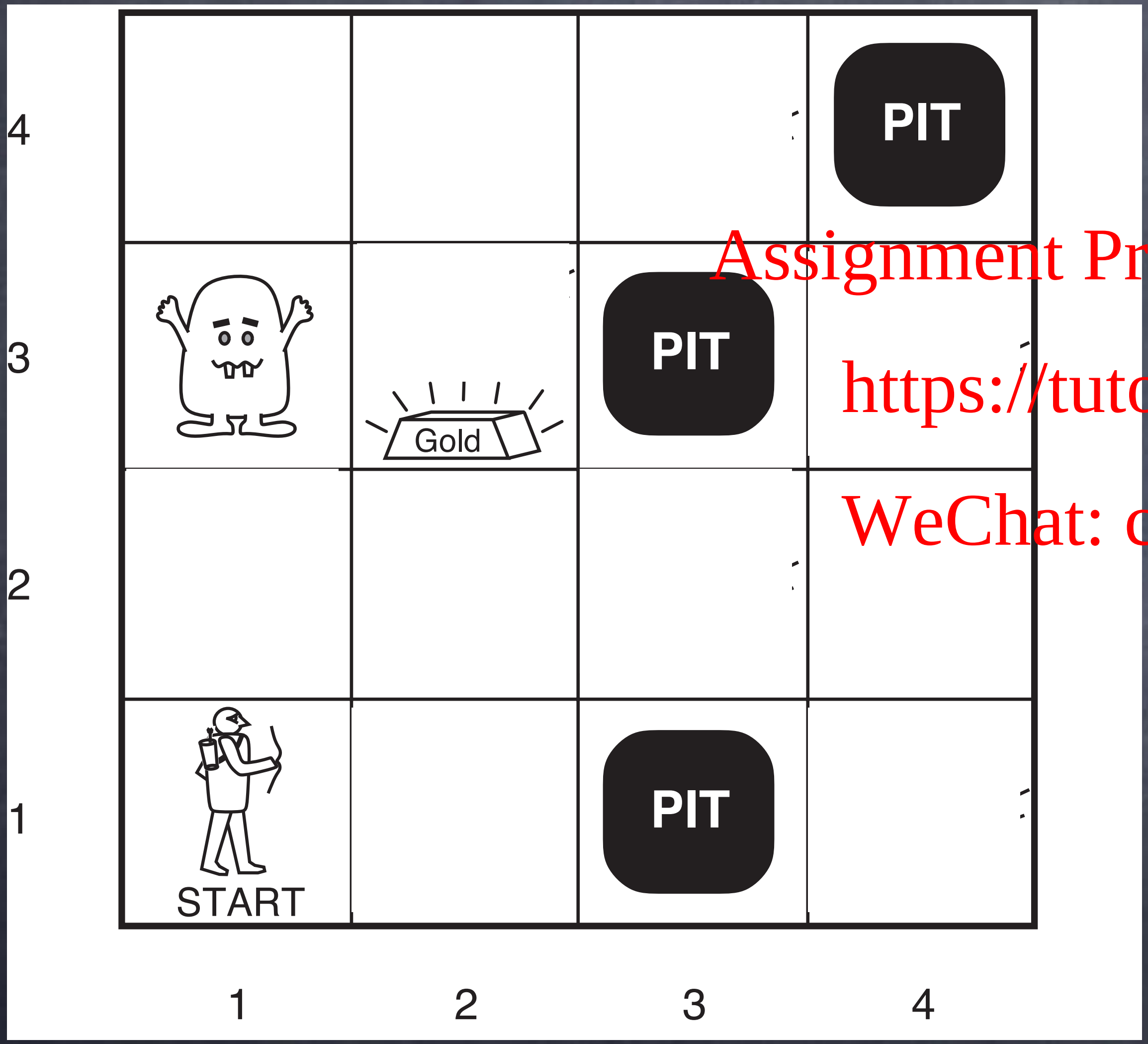
Hunt The Wumpus



Hunt The Wumpus

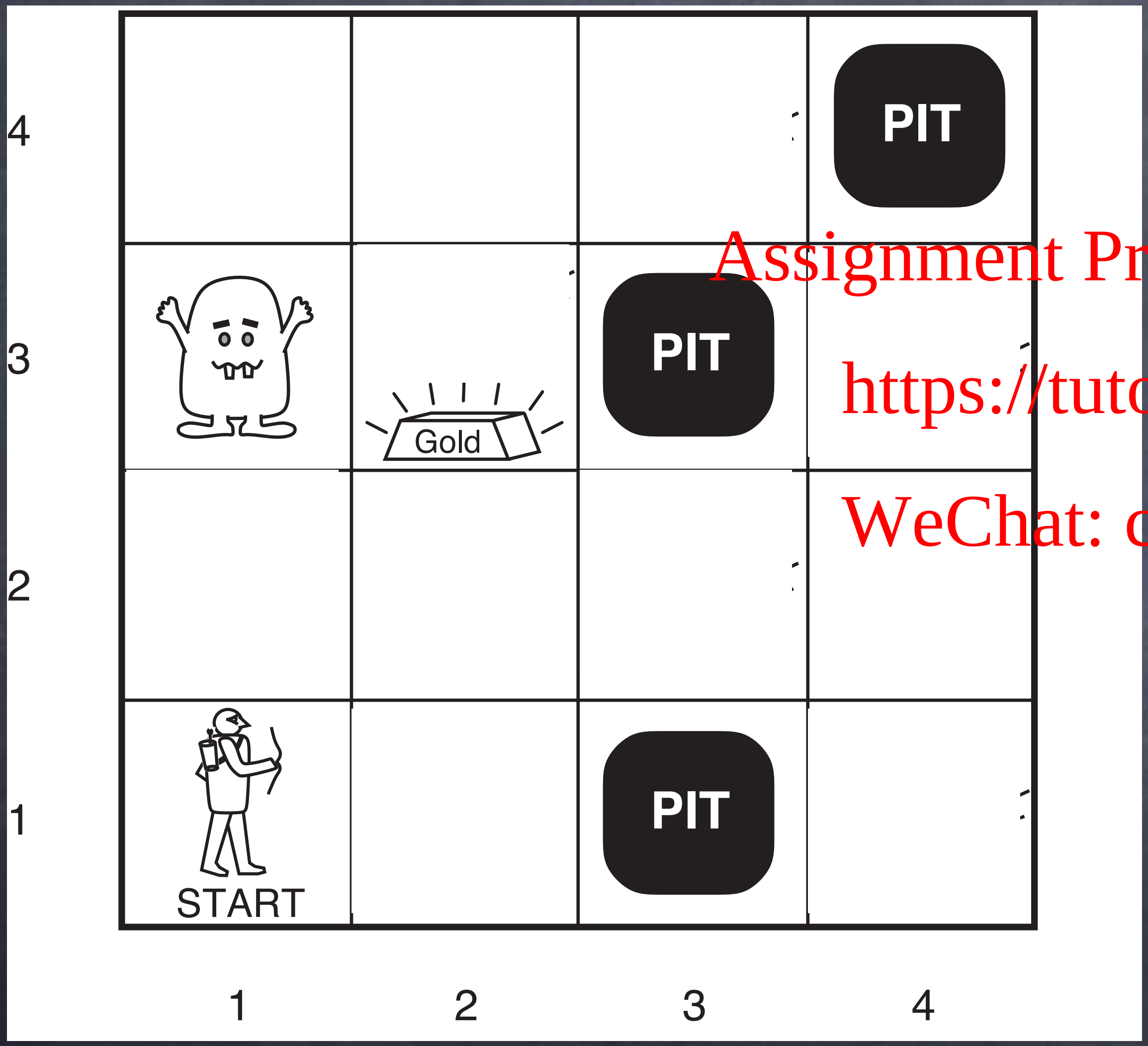


Hunt The Wumpus



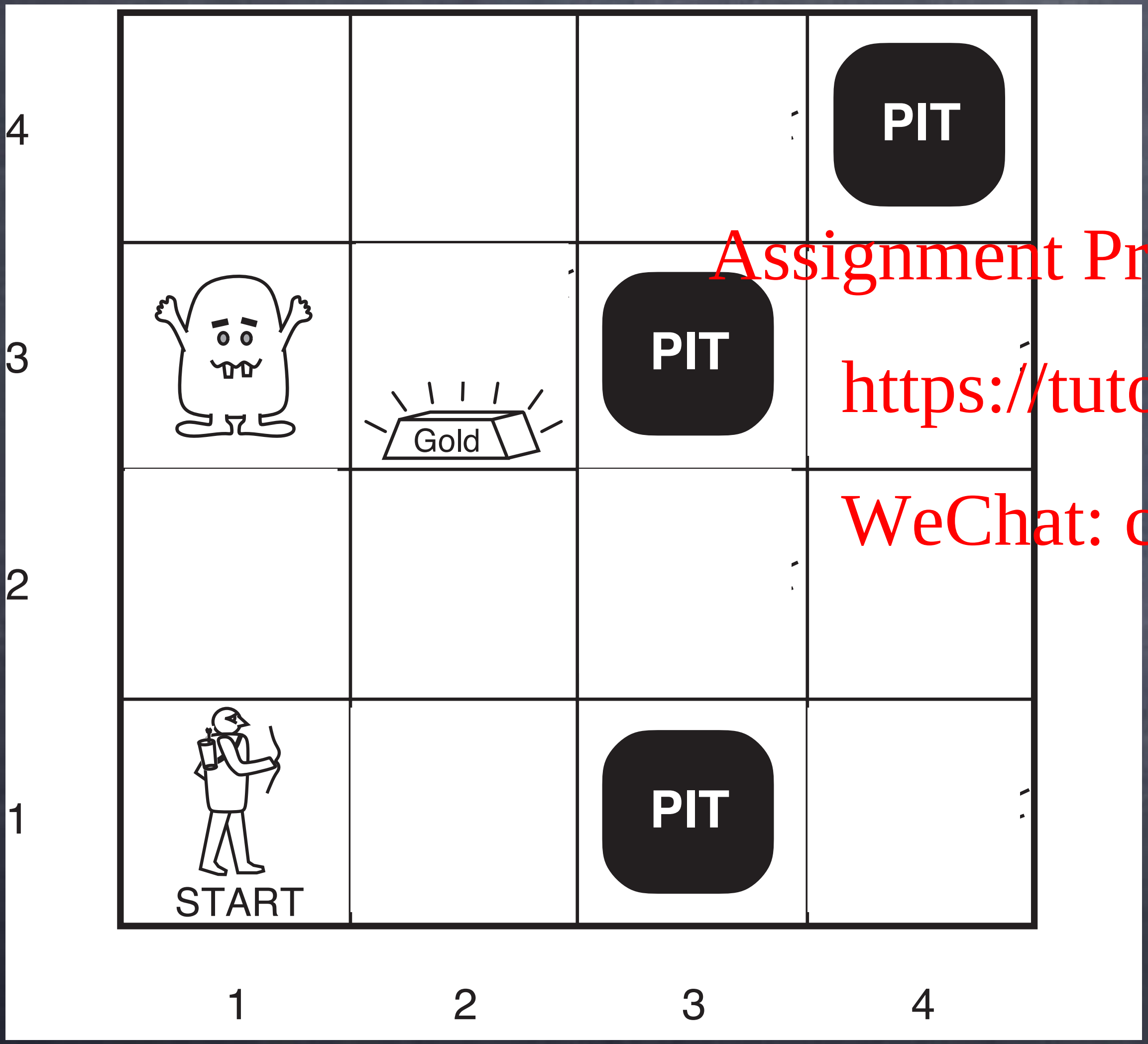
- Move fwd, turn left/ right
- Shoot arrow (once)
- Grab gold
- Climb out (from [1,1])
- Costs:
  - -1 per move
  - -10 to shoot

Hunt The Wumpus

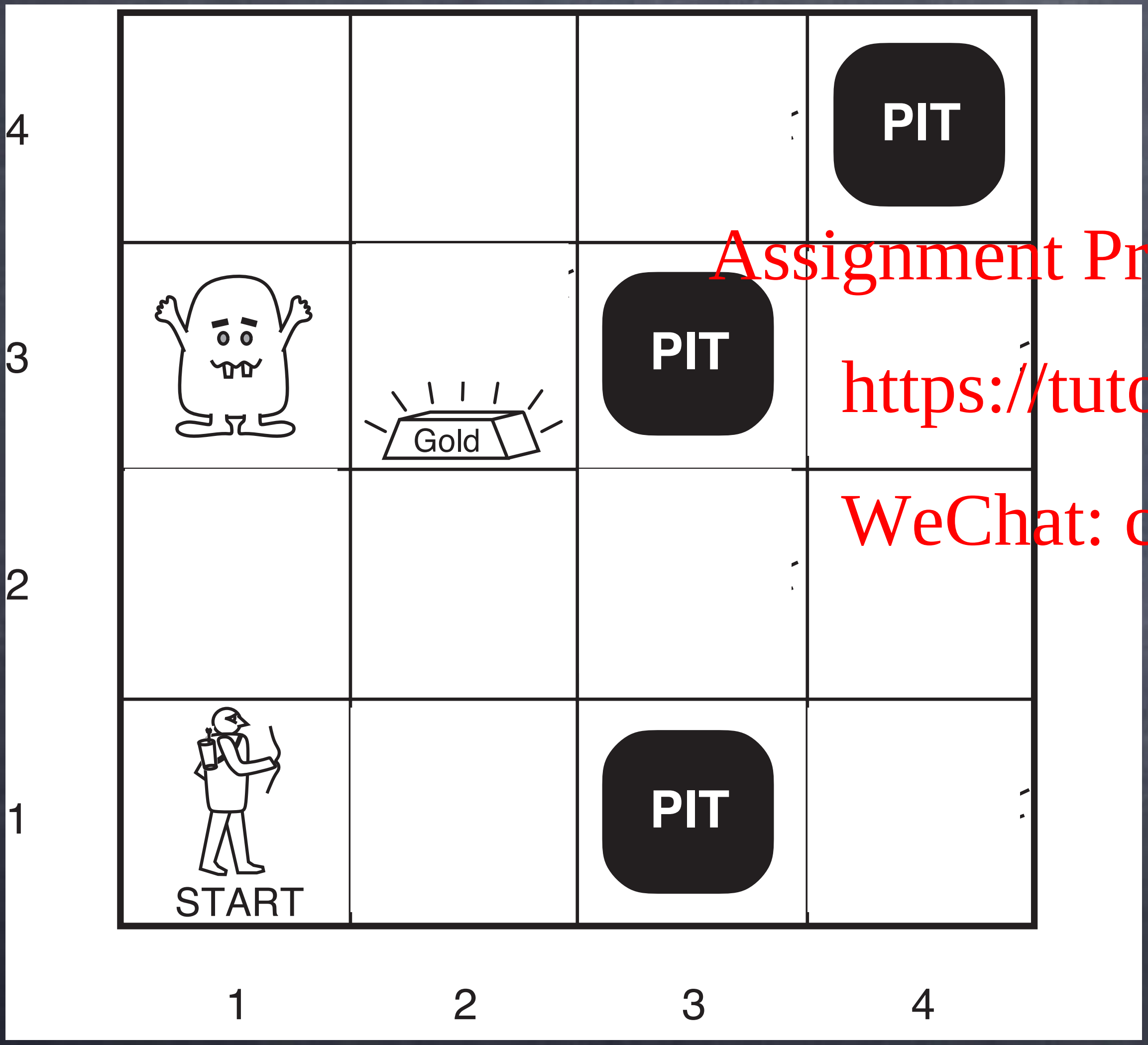


- Get gold and climb out: +1000
- Fall in pit or get eaten by wumpus: -1000

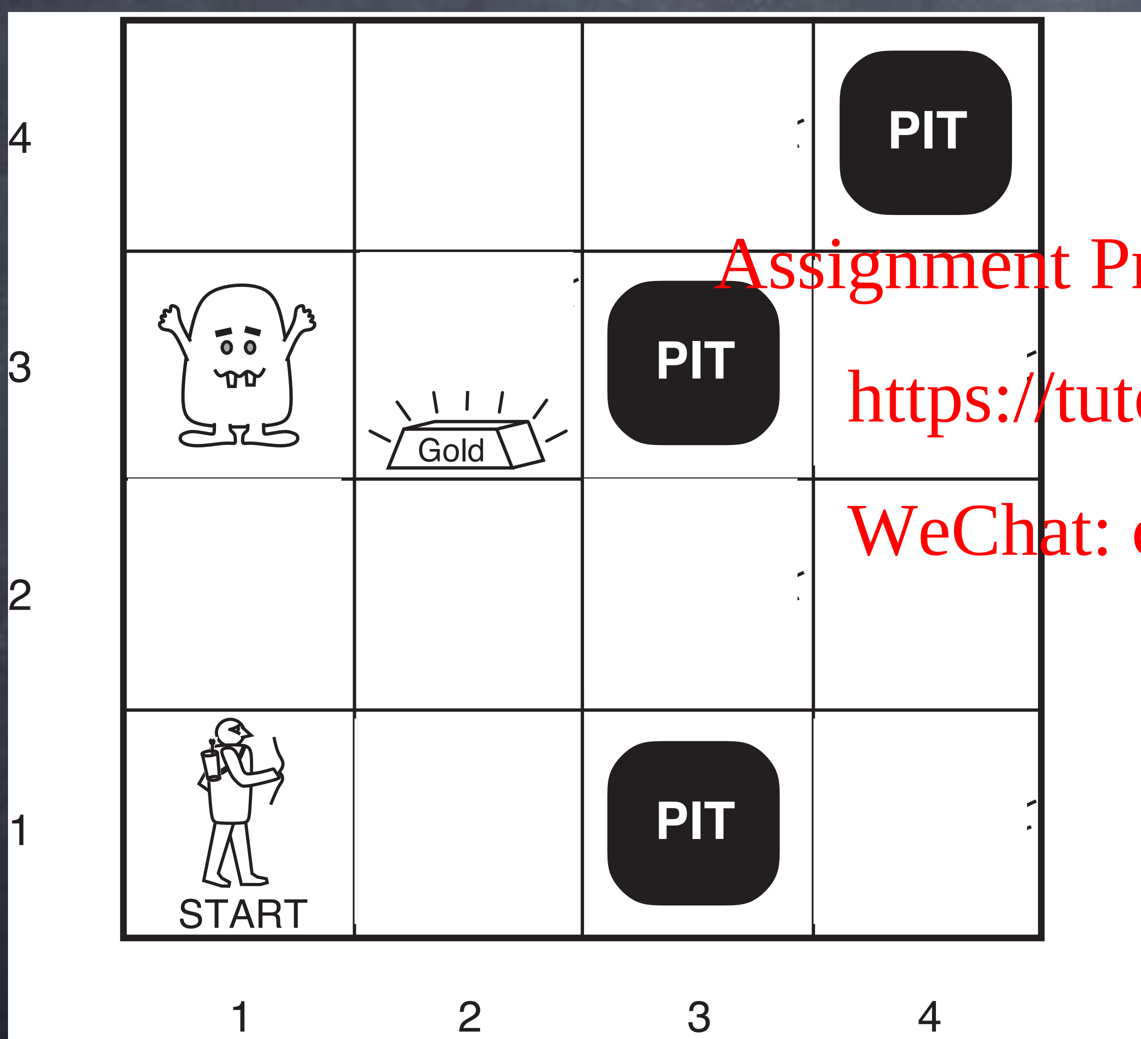
Hunt The Wumpus

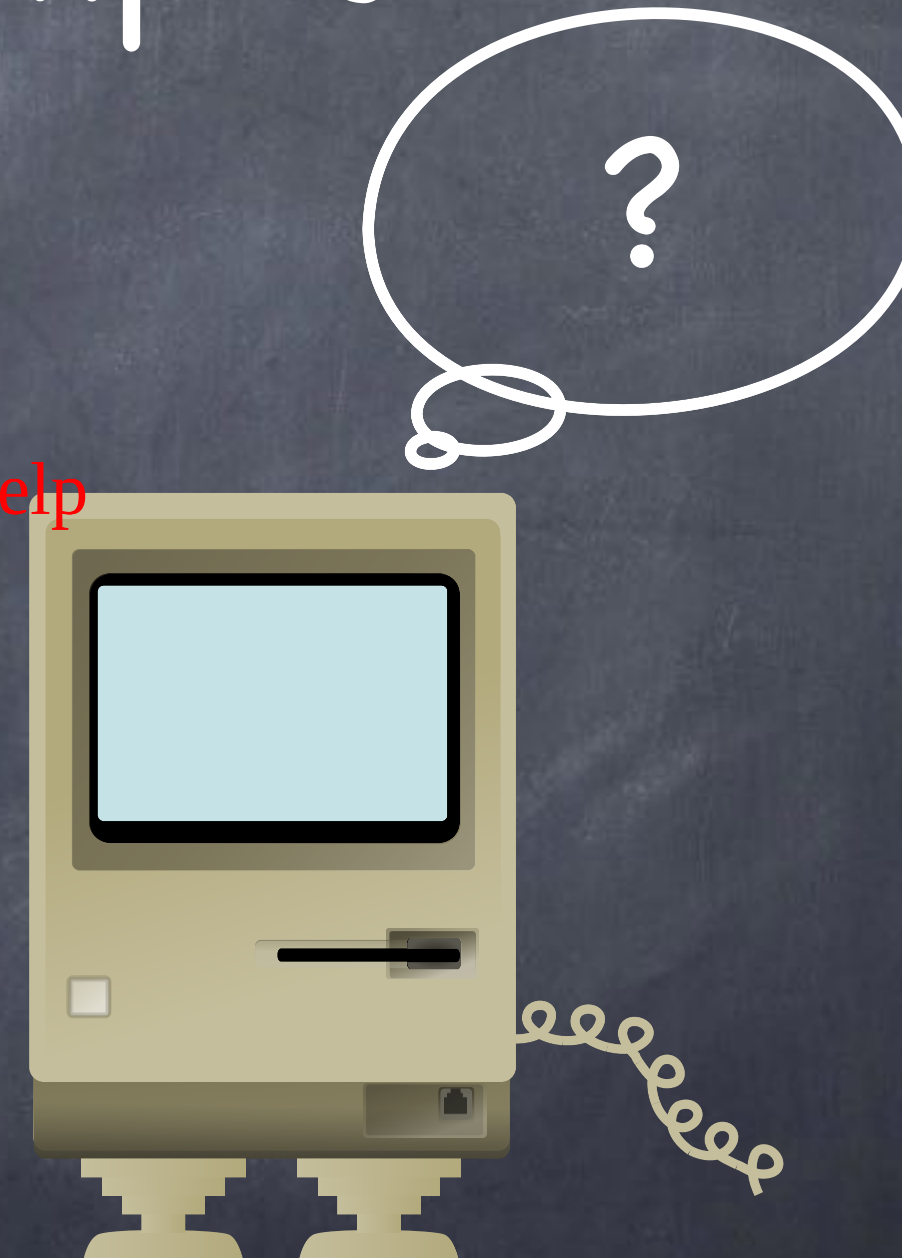


Hunt The Wumpus



Hunt The Wumpus

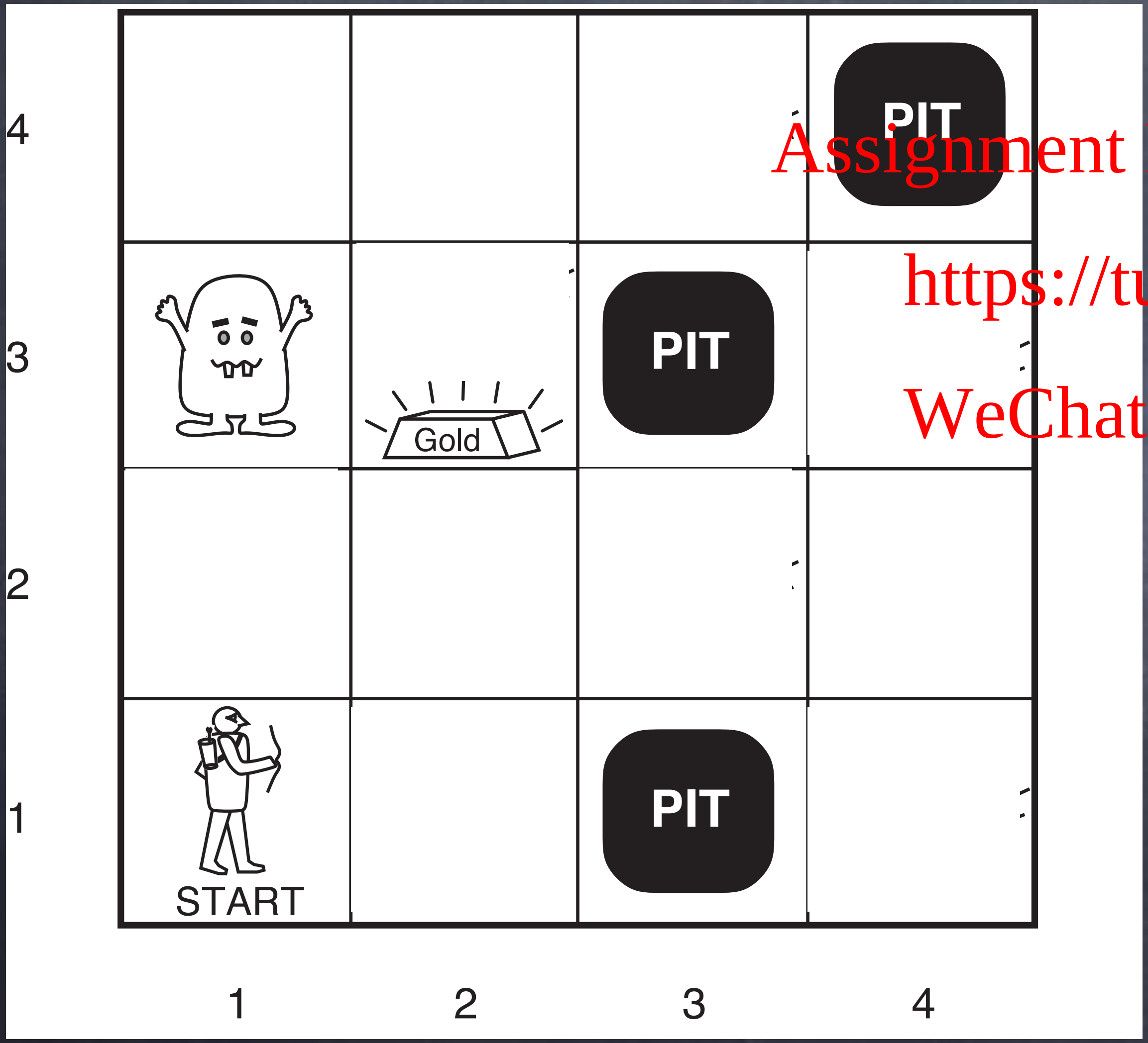




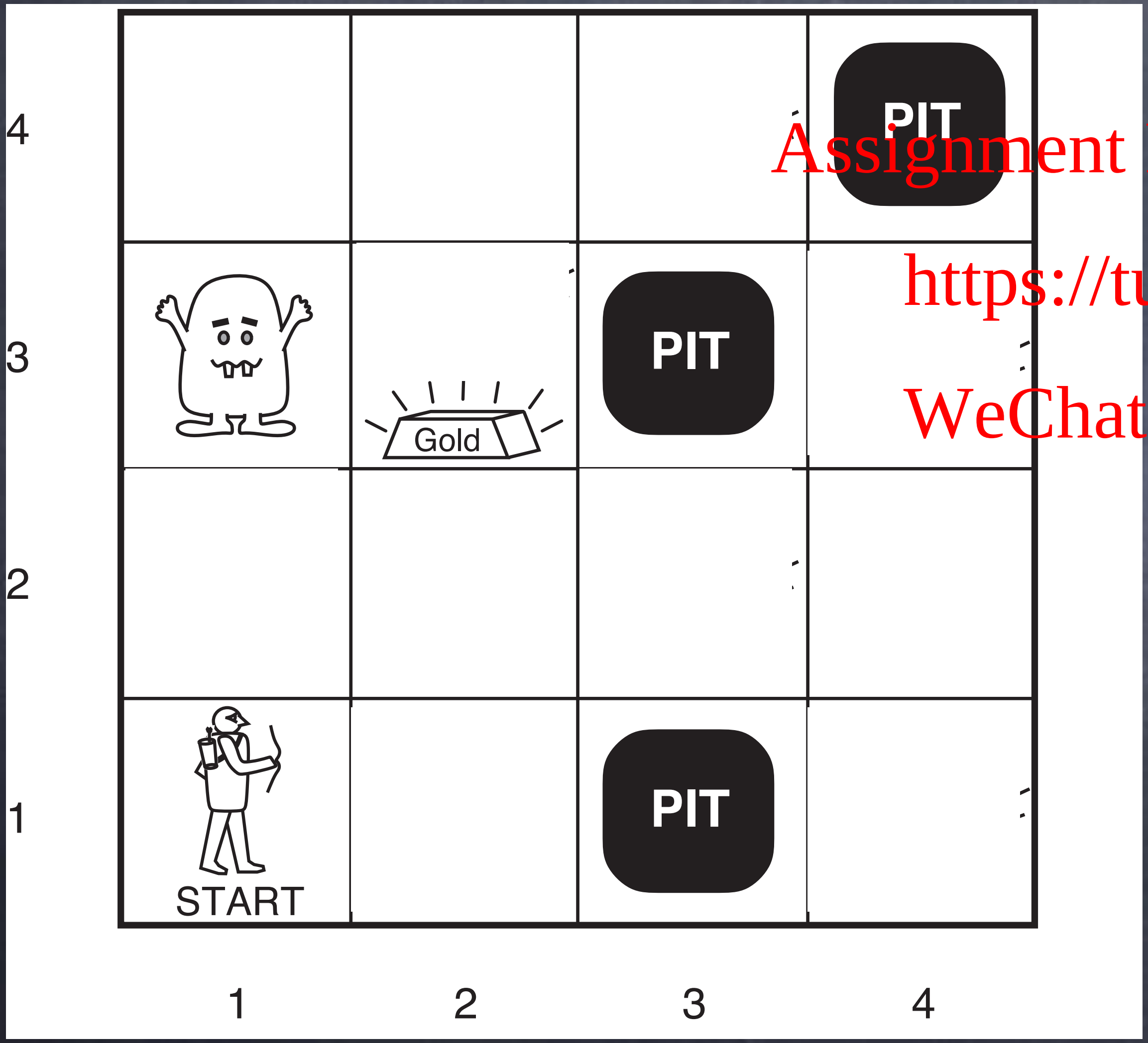
Boolean CSP

- All variables must be Booleans
  - Domains all $\{ \, true, \, false \, \}$

WW Boolean CSP



WW Boolean CSP



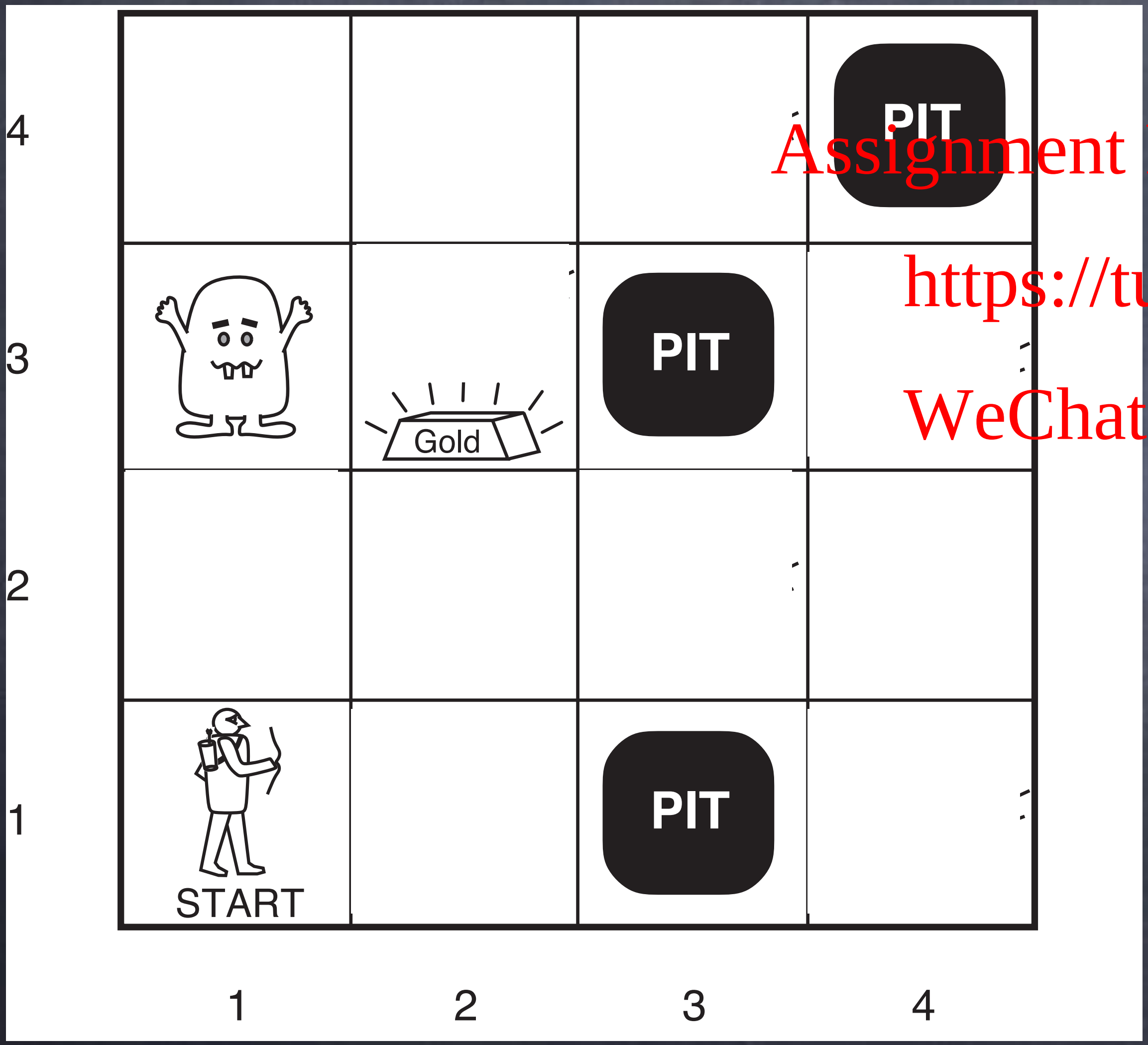
For each room:

Has pit? true, false, idk

Has wumpus? true, false, idk

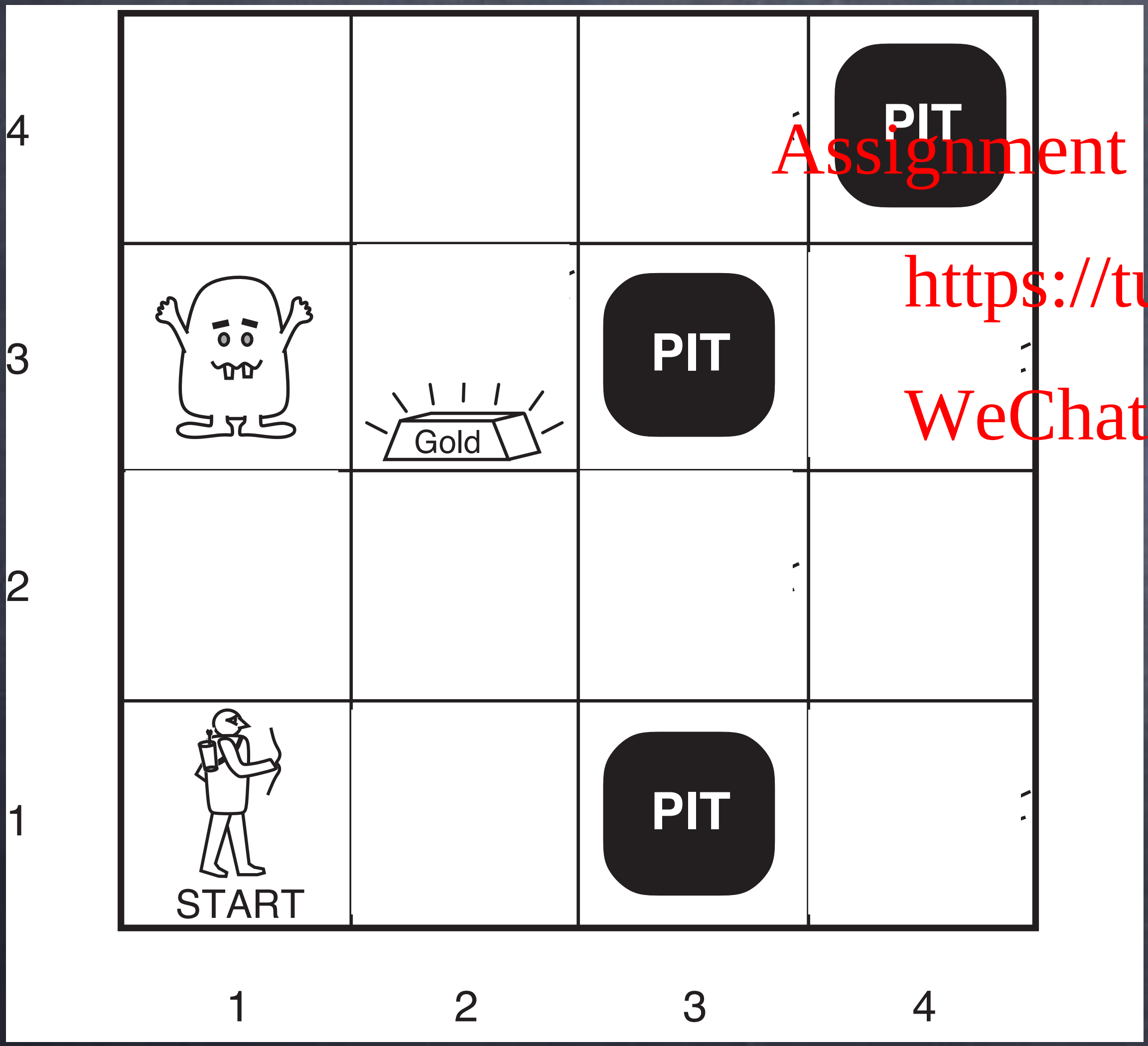
Has gold? true, false, idk

WW Boolean CSP



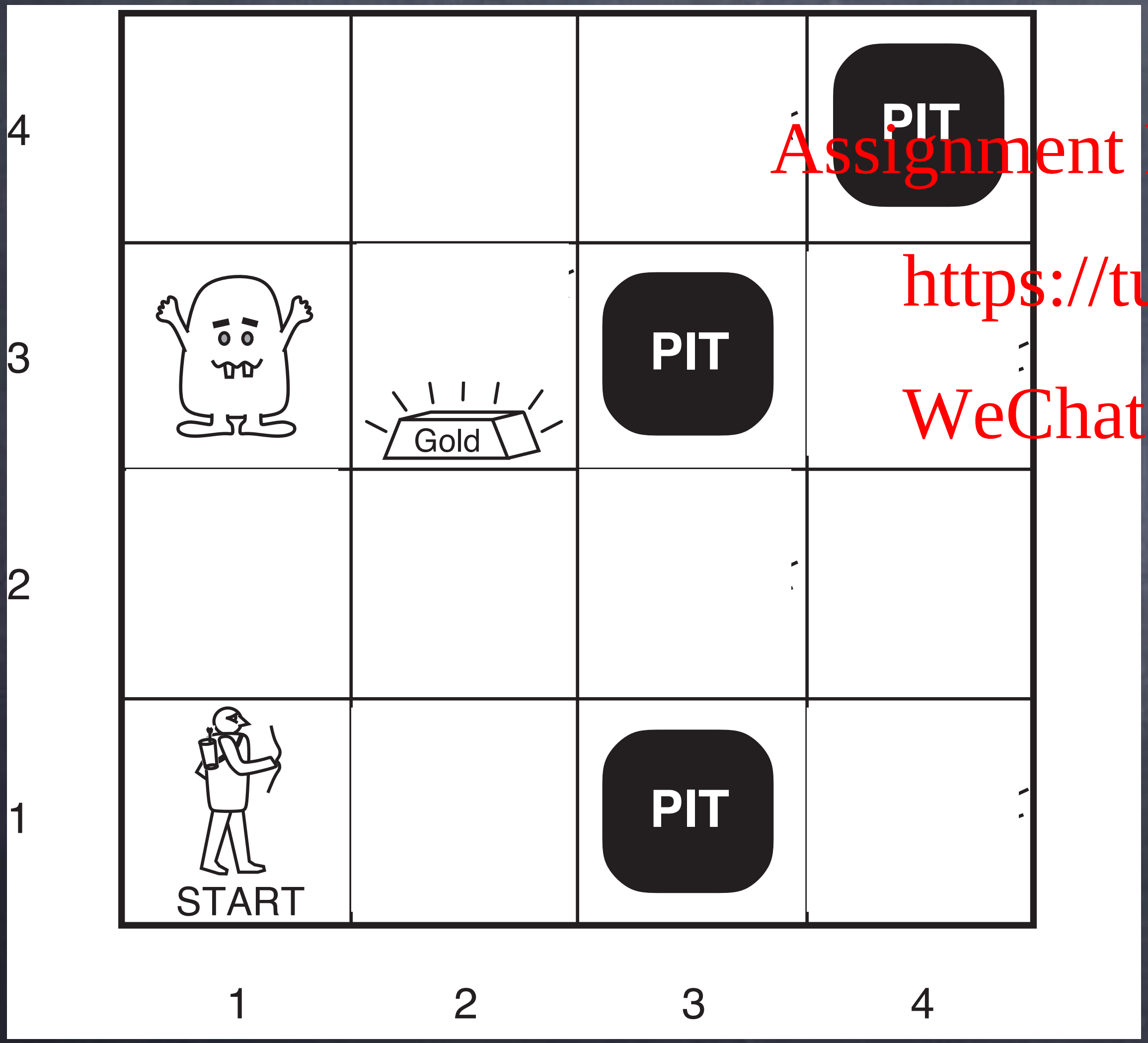

```
Boolean[][] P =
    new Boolean[n][n];
Boolean[][] W =
    new Boolean[n][n];
Boolean[][] G =
    new Boolean[n][n];

// P[i][j] == Boolean.TRUE
//     if there's a pit at [i,j]
// P[i][j] == Boolean.FALSE
//     if there's no pit at [i,j]
// P[i][j] == null
//     if we don't know
// W[i][j] ditto for wumpus
// G[i][j] ditto for gold
```

WW Boolean CSP



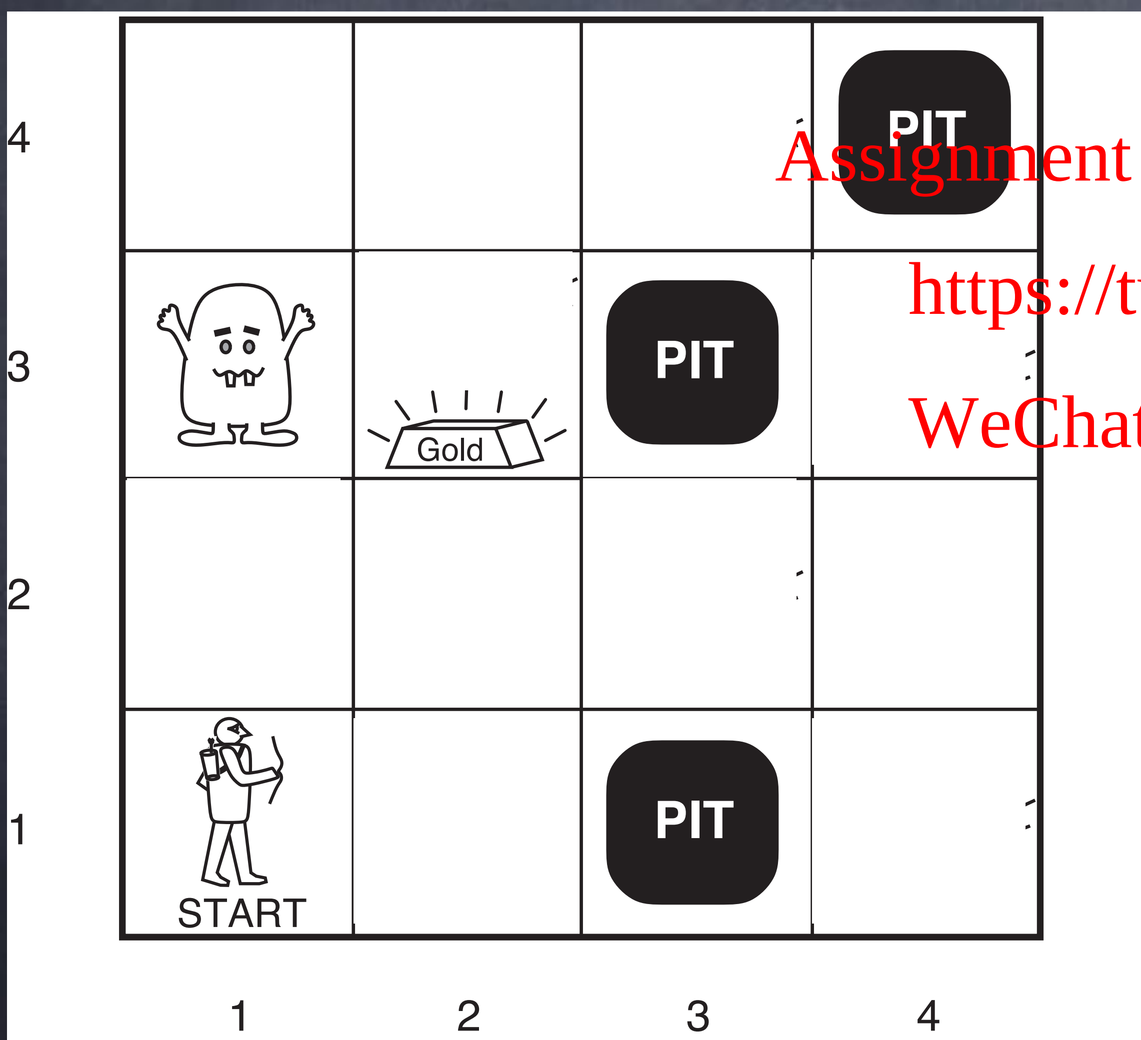

```
Boolean P_11, P_12, P_13, ...,
        P_21, P_22, ...;
Boolean W_11, W_12, W_13, ...,
        W_21, W_22, ...;
Boolean G_11, G_12, G_13, ...,
        G_21, G_22, ...;

// P_ij == Boolean.TRUE
//    if there's a pit at [i,j]
// P_ij == Boolean.FALSE
//    if there's no pit at [i,j]
// P_ij == null
//    if we don't know
// W_ij ditto for wumpus
// G_ij ditto for gold
```

WW Boolean CSP



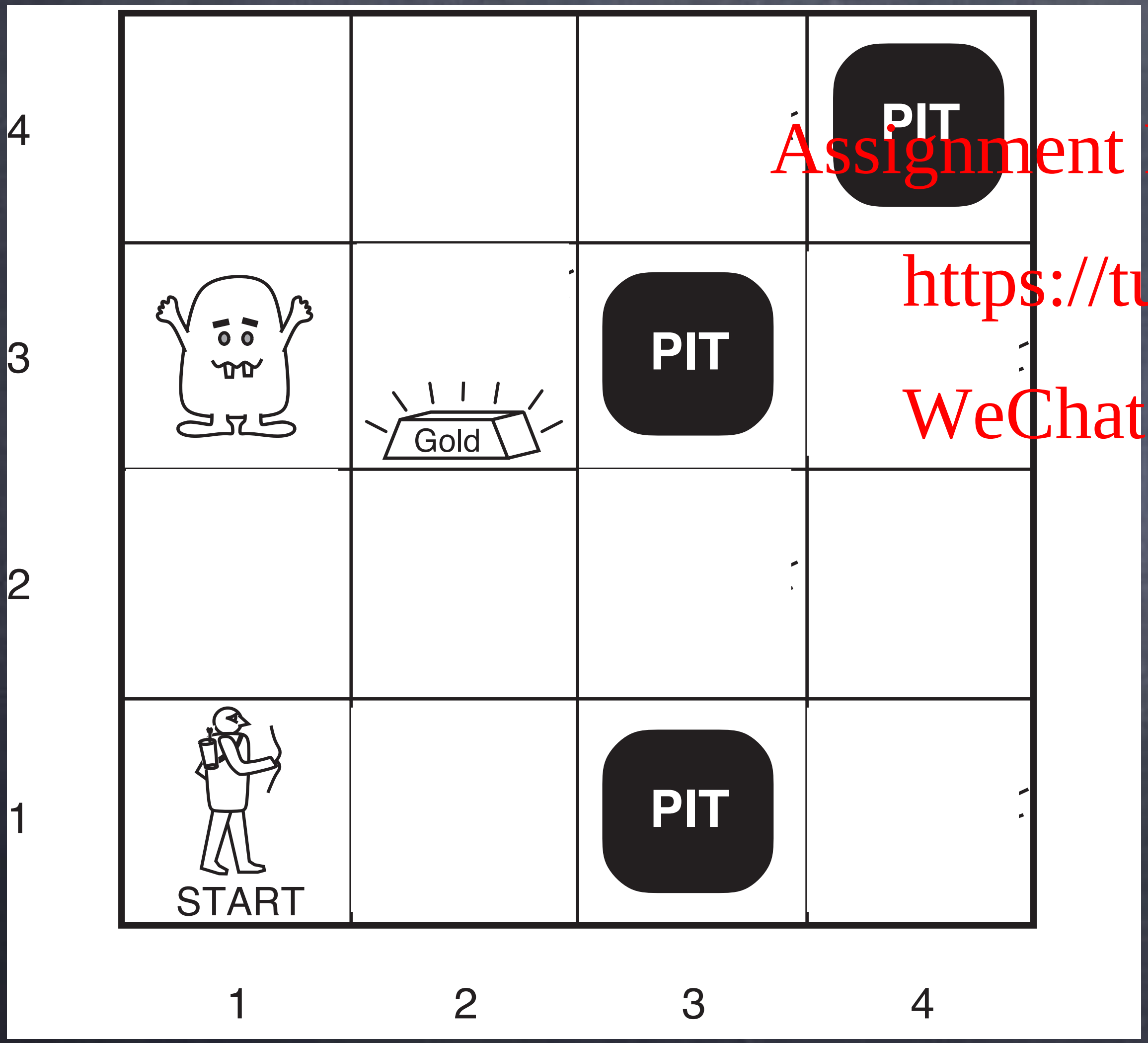
Position pos = <i, j>

WW Boolean CSP



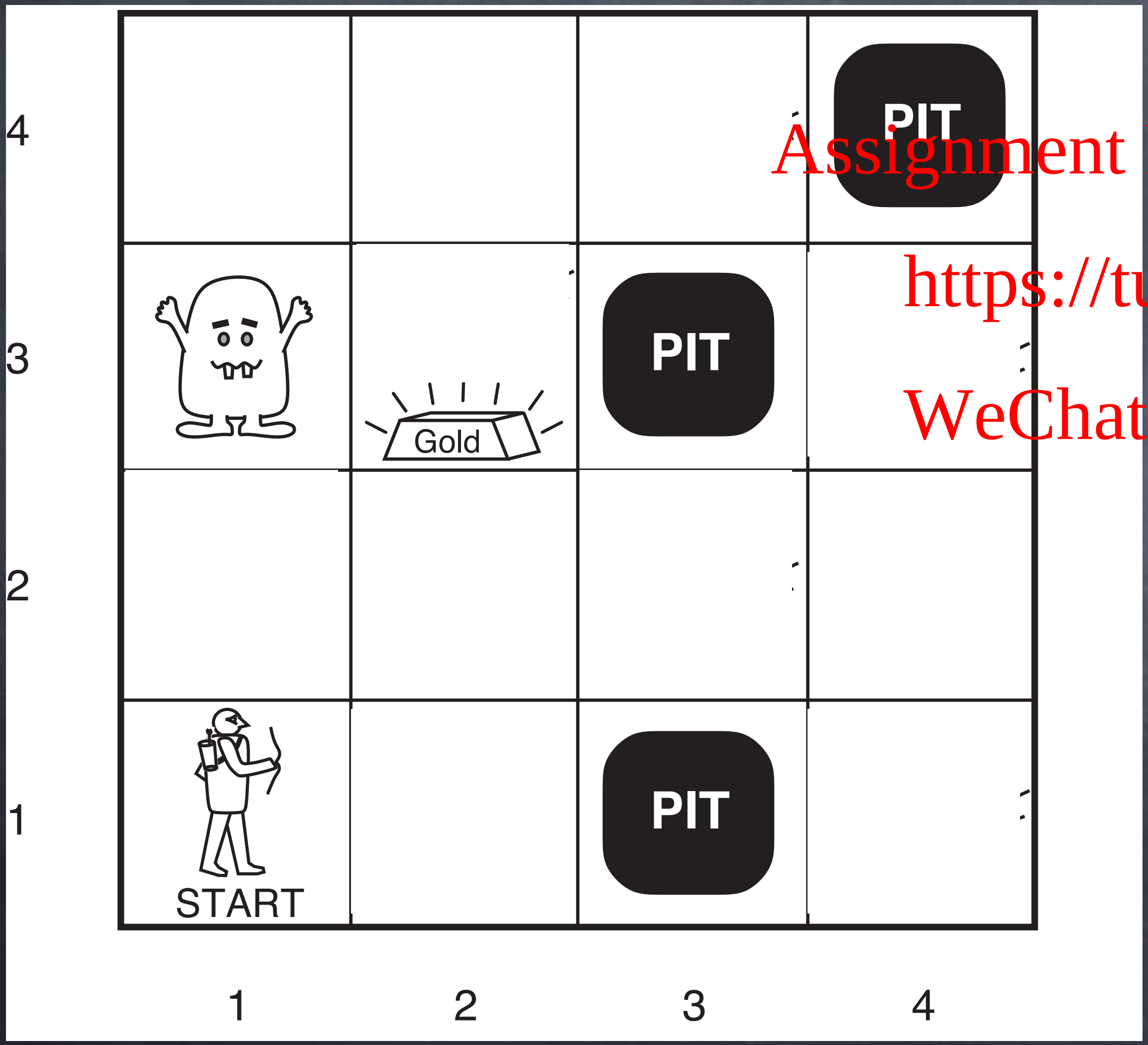

```
Boolean L_11 = true,
        L_12 = false,
        L_13 = false,
        ...;

// L_ij == Boolean.TRUE
//   if agent is at location [i,j]
//   else Boolean.FALSE
// Note: The agent always knows
//         where it is.
```

WW Boolean CSP



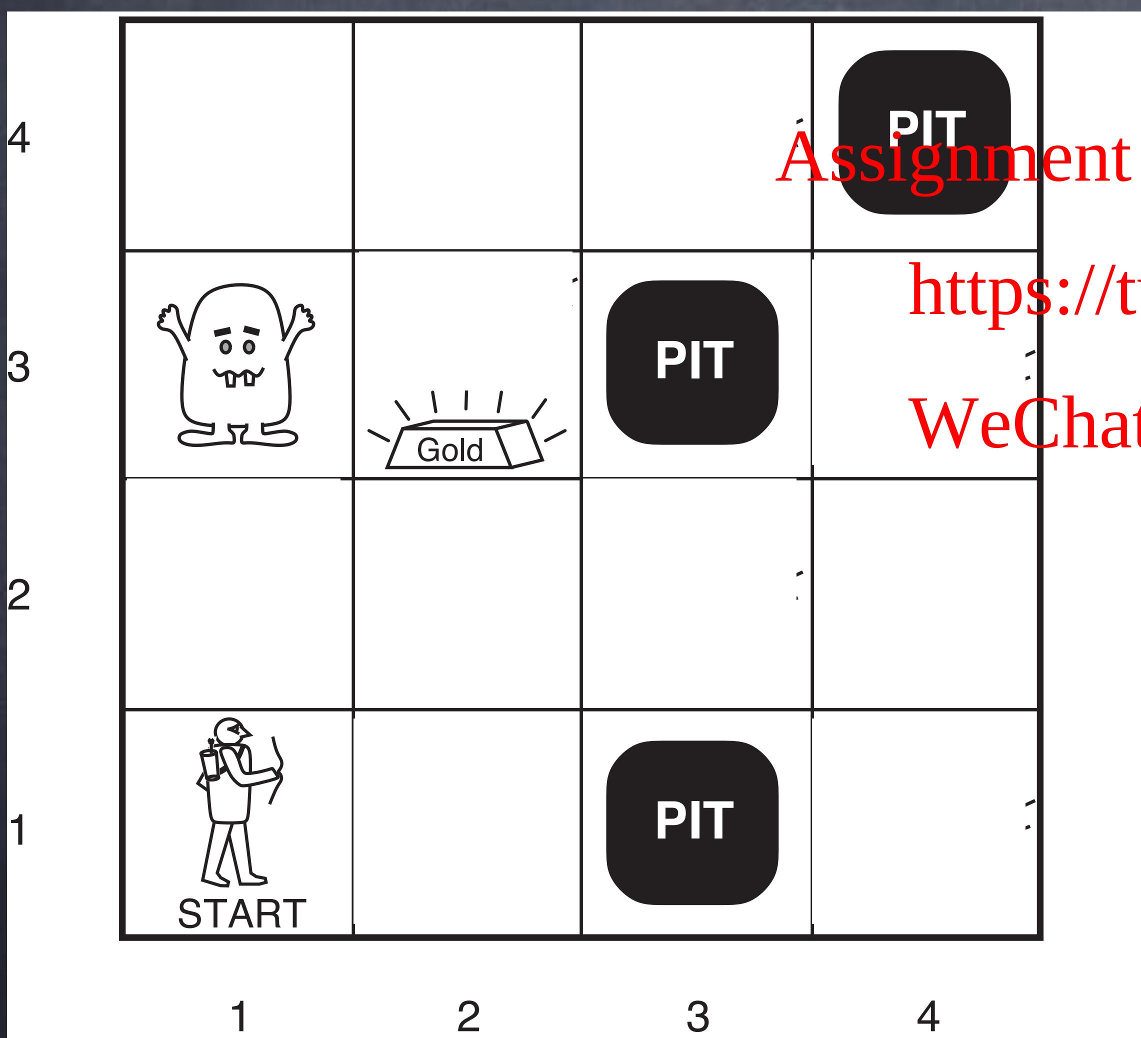

```
enum Orientation = { N,S,E,W };

Orientation orientation;
```

WW Boolean CSP



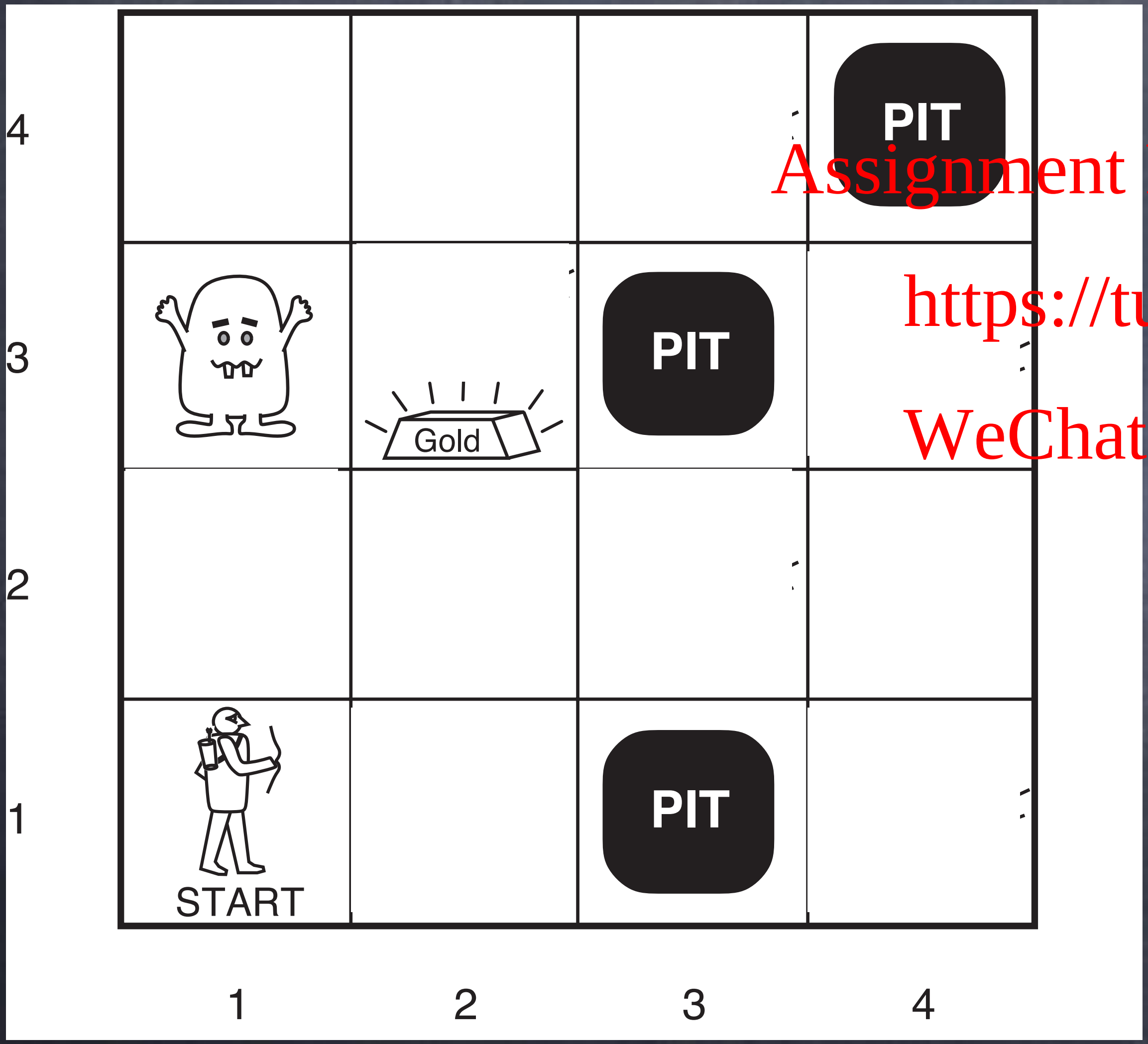

```
Boolean Facing_N, Facing_S,
        Facing_E, Facing_W;

// FacingN == Boolean.TRUE
    if agent is facing north
// Ditto for the other directions
// Note: The agent always knows
//       which way it is facing
```

WW Boolean CSP

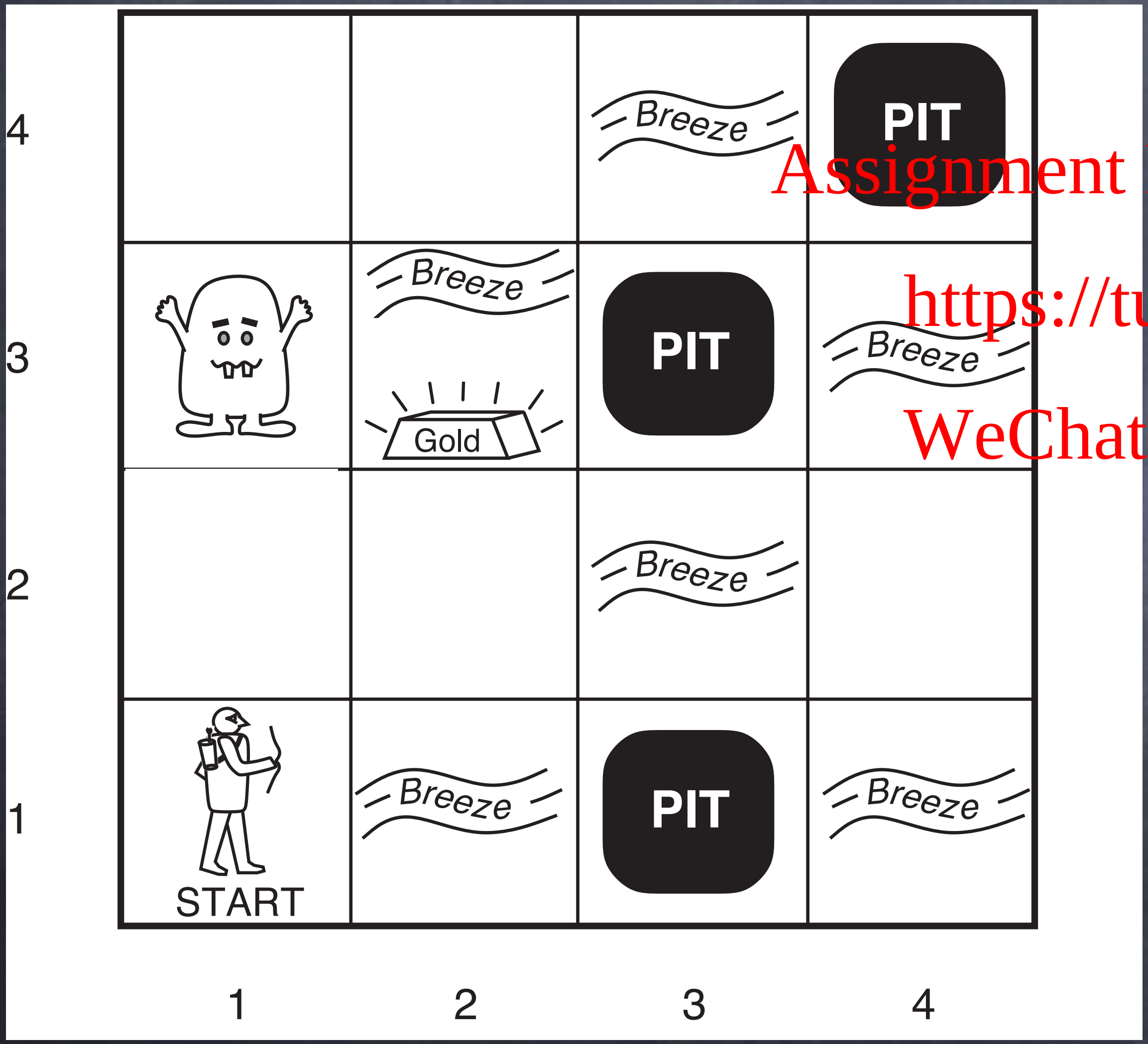



```
Boolean P_11, P_12, P_13, ...,
        P_21, P_22, ...;
Boolean W_11, W_12, W_13, ...,
        W_21, W_22, ...;
Boolean G_11, G_12, G_13, ...,
        G_21, G_22, ...;
Boolean L_11 = true,
        L_12 = false,
        L_13 = false,
        ...;
Boolean Facing_N, Facing_S,
        Facing_E, Facing_W;
```

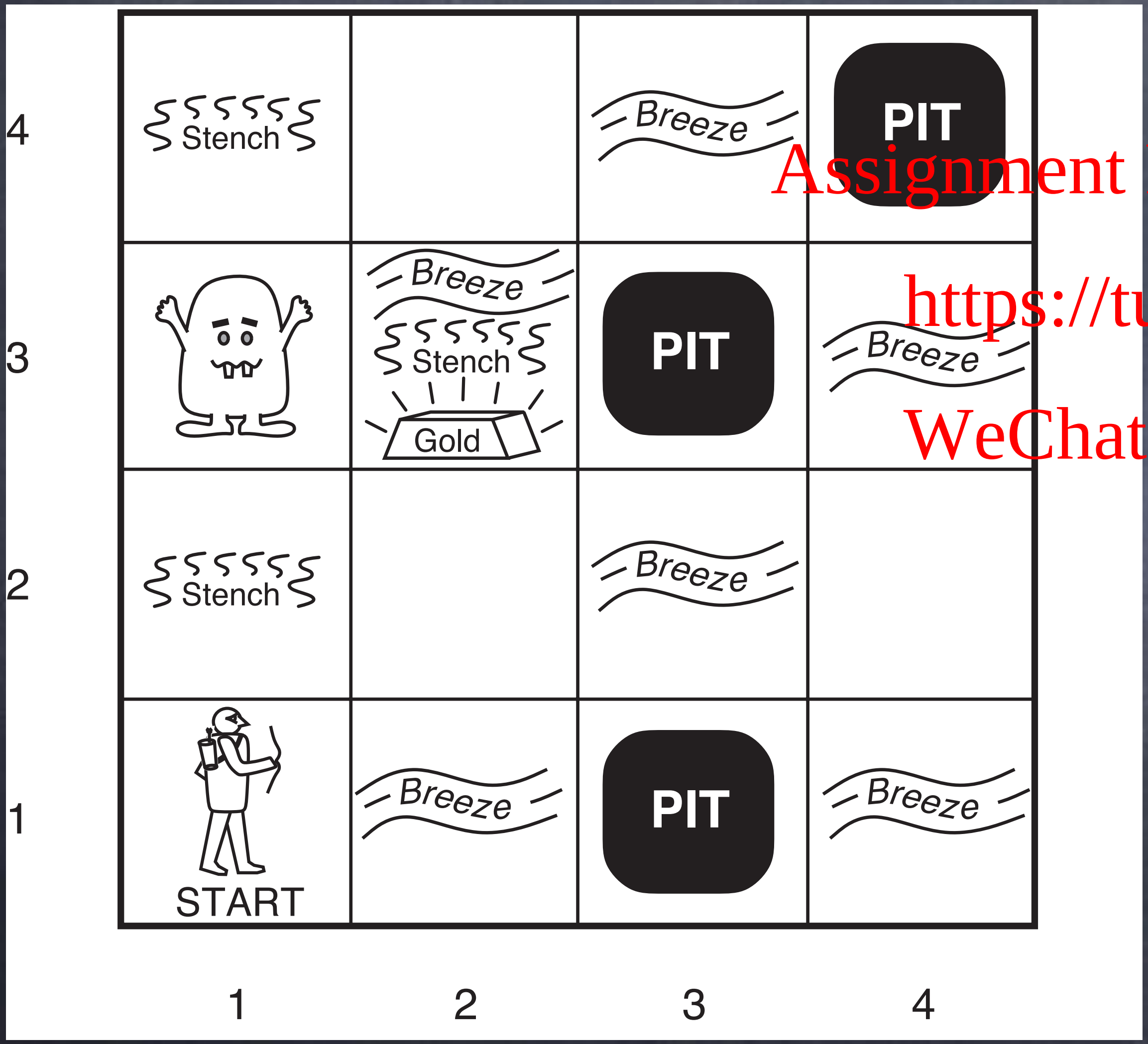
Hunt The Wumpus



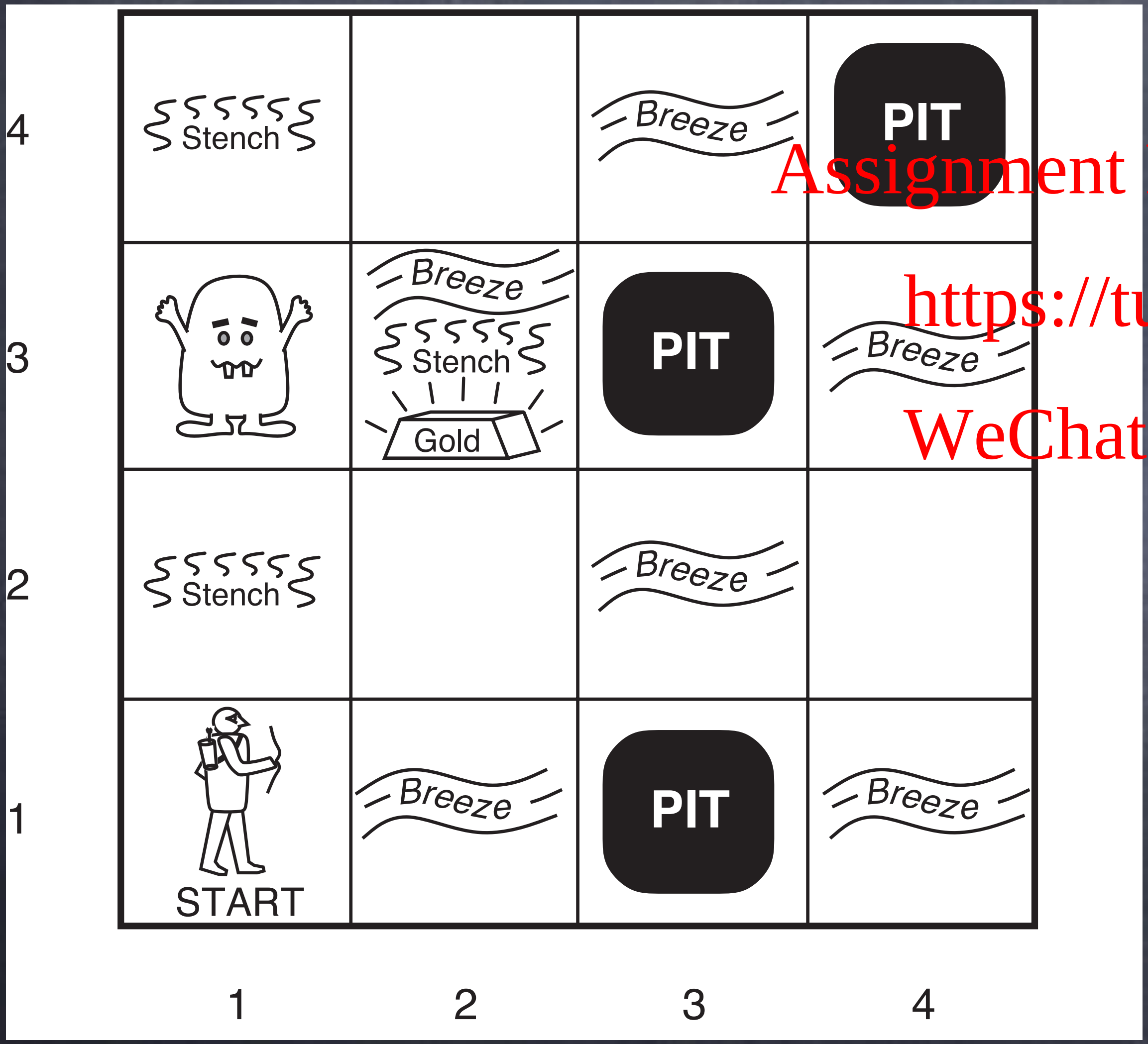
Hunt The Wumpus



Hunt The Wumpus

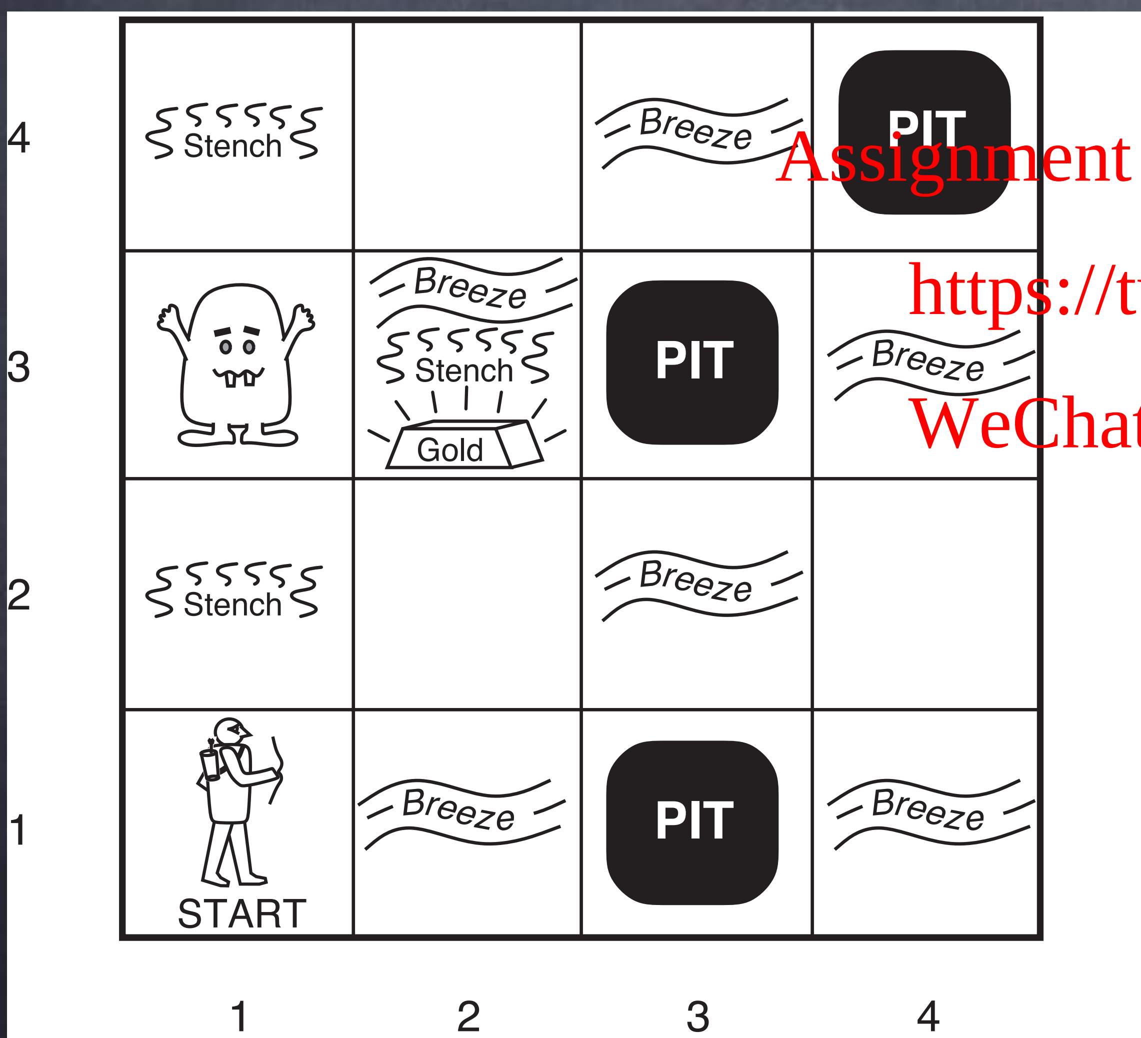


Hunt The Wumpus



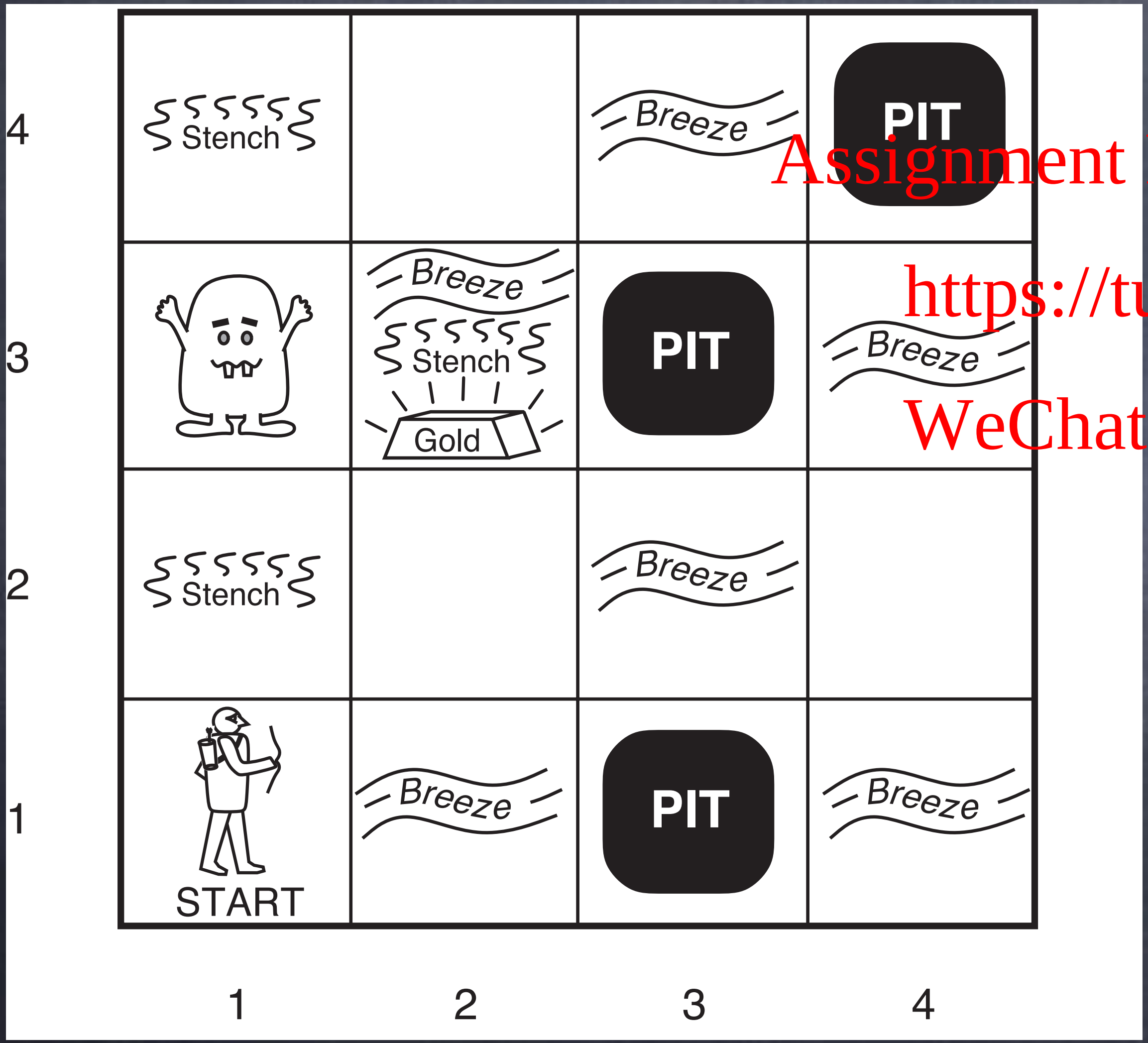
- Breeze
- Stench
- Glimmer
- Bump
- Scream

WW Boolean CSP




```
Boolean S_11, S_12, S_13, ...,
        S_21, S_22, ...;
Boolean B_11, B_12, B_13, ...,
        B_21, B_22, ...;
...

// S_ij == Boolean.TRUE
//    if a stench perceived @[i,j]
// S_ij == Boolean.FALSE
//    if no stench @[i,j]
// S_ij == null
//    if we don't know
// B_ij ditto for breeze
// Ditto for other percepts
```

WW Boolean CSP




```
Boolean P_11, P_12, P_13, ...,
        P_21, P_22, ...;
Boolean W_11, W_12, W_13, ...,
        W_21, W_22, ...;
Boolean G_11, G_12, G_13, ...,
        G_21, G_22, ...;
Boolean L_11 = true,
        L_12 = false,
        L_13 = false,
        ...;
Boolean Facing_N, Facing_S,
        Facing_E, Facing_W;
Boolean S_11, S_12, S_13, ...,
        S_21, S_22, ...;
Boolean B_11, B_12, B_13, ...,
        B_21, B_22, ...;
...
```

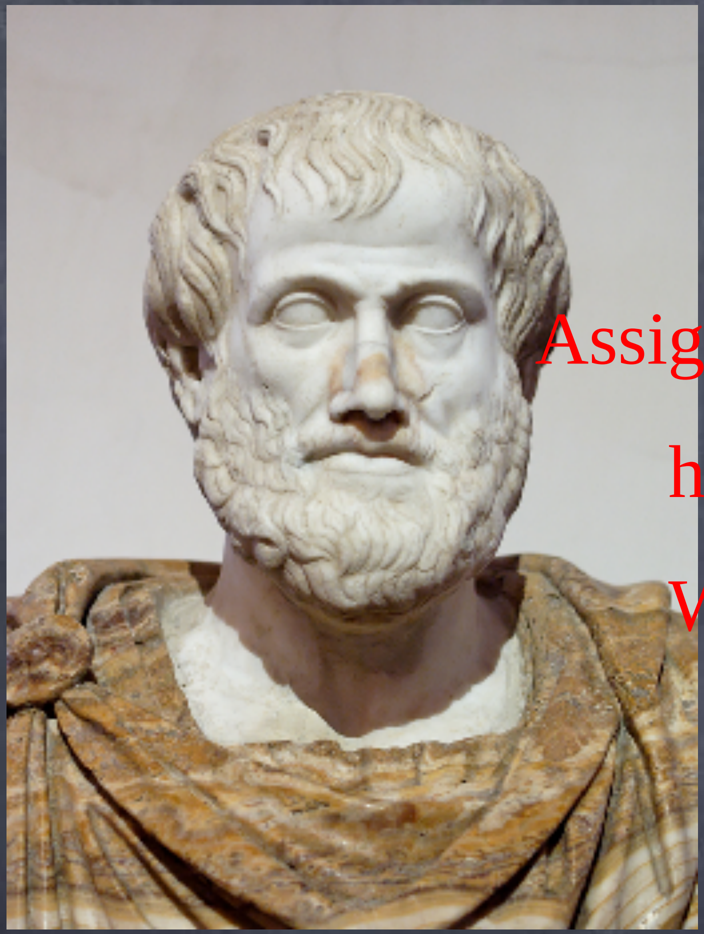
Constraints

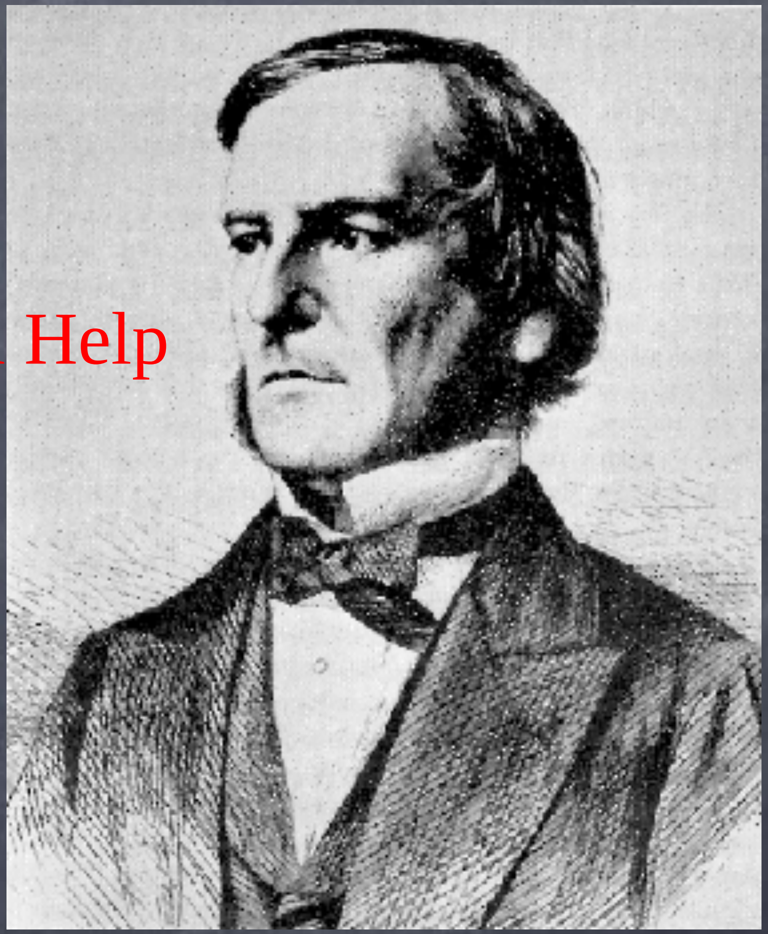
Constraints

- Constraints are Boolean functions of the variables
    - $WA \neq NT$
    - $T_2 \geq T_1 + 10$
    - $AllDiff(A1, A2, A3, B1, B2, B3, C1, C2, C3)$

Constraints

- Constraints are Boolean functions of the variables
  - $WA \neq NT$
  - $T_2 \geq T_1 + 10$
  - $AllDiff(A1, A2, A3, B1, B2, B3, C1, C2, C3)$
- Constraints on Boolean variables are Boolean functions of Boolean values





Aristole

(384BC - 332BC)

George Boole

(1815-1864)

Propositional Logic (Boolean Algebra)

A Programming Language for Knowledge!

Propositions

- Propositions: things that can be true or false

Propositions

- Propositions: things that can be true or false

- Atomic propositions:
  - "It is raining"
  - "Socrates was a person"
  - "The wumpus is in room [2,2]"

Connectives (Operators)

- Combine propositions into larger propositions

  - $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$

  - $\neg Raining$
  - $Raining \wedge Cold$, $Raining \vee Sunny$
  - $Raining \wedge BelowFreezing \Rightarrow Slippery$

Connectives (Operators)

- Combine propositions into larger propositions

Syntax

  - $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$

  - $\neg Raining$
  - $Raining \wedge Cold$, $Raining \vee Sunny$
  - $Raining \wedge BelowFreezing \Rightarrow Slippery$

Connectives (Operators)

- Combine propositions into larger propositions
  - $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$
- Each connective represents a Boolean function of its (Boolean) arguments

Syntax

Semantics

Connectives

p		$\neg p$				
F		T				
T		F				

Connectives

p	q	$\neg p$	$p \wedge q$	$p \vee q$	$p \Rightarrow q$	$p \Leftrightarrow q$
F	F	T	F	F	T	T
F	T	F	F	T	T	F
T	F		F	T	F	F
T	T		T	T	T	T

Sentences (Expressions)

- Every sentence of propositional logic represents a Boolean function of its (Boolean) arguments

Sentences (Expressions)

- Every sentence of propositional logic represents a Boolean function of its (Boolean) arguments
- The meaning of a sentence (the Boolean function that it denotes) is the composition of its parts

Sentences

$L_{1,1}$: True if the agent is in room [1,1]

$W_{1,2}$: True if the wumpus is in room [1,2]

$W_{2,1}$: True if the wumpus is in room [2,1]

Sentences

$L_{1,1}$	$W_{1,2}$	$W_{2,1}$
F	F	F
F	F	T
F	T	F
F	T	T
T	F	F
T	F	T
T	T	F
T	T	T

Sentences

$$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$$

$L_{1,1}$	$W_{1,2}$	$W_{2,1}$
F	F	F
F	F	T
F	T	F
F	T	T
T	F	F
T	F	T
T	T	F
T	T	T

Sentences

$$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$$

$L_{1,1}$	$W_{1,2}$	$W_{2,1}$	$W_{1,2} \vee W_{2,1}$	$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$
F	F	F	F	F
F	F	T	T	F
F	T	F	T	F
F	T	T	T	F
T	F	F	F	F
T	F	T	T	T
T	T	F	T	T
T	T	T	T	T

Truth Table

$$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$$

$L_{1,1}$	$W_{1,2}$	$W_{2,1}$	$W_{1,2} \vee W_{2,1}$	$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$
F	F	F	F	F
F	F	T	T	F
F	T	F	T	F
F	T	T	T	F
T	F	F	F	F
T	F	T	T	T
T	T	F	T	T
T	T	T	T	T

Propositional Logic and Boolean CSPs

Sentences

$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$

$L_{1,1}$	$W_{1,2}$	$W_{2,1}$	$W_{1,2} \vee W_{2,1}$	$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$
F	F	F	F	F
F	F	T	T	F
F	T	F	T	F
F	T	T	T	F
T	F	F	F	F
T	F	T	T	T
T	T	F	T	T
T	T	T	T	T

Possible Worlds

$$L_{1,1} \land (W_{1,2} \lor W_{2,1})$$

$L_{1,1}$	$W_{1,2}$	$W_{2,1}$	$W_{1,2} \lor W_{2,1}$	$L_{1,1} \land (W_{1,2} \lor W_{2,1})$
F	F	F	F	F
F	F	T	T	F
F	T	F	T	F
F	T	T	T	F
T	F	F	F	F
T	F	T	T	T
T	T	F	T	T
T	T	T	T	T

Sentences

$$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$$

$L_{1,1}$	$W_{1,2}$	$W_{2,1}$	$W_{1,2} \vee W_{2,1}$	$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$
F	F	F	F	F
F	F	T	T	F
F	T	F	T	F
F	T	T	T	F
T	F	F	F	F
T	F	T	T	T
T	T	F	T	T
T	T	T	T	T

Satisfiability

$$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$$

$L_{1,1}$	$W_{1,2}$	$W_{2,1}$	$W_{1,2} \vee W_{2,1}$	$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$
F	F	F	F	F
F	F	T	T	F
F	T	F	T	F
F	T	T	T	F
T	F	F	F	F
T	F	T	T	T
T	T	F	T	T
T	T	T	T	T

Impossible Worlds

$$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$$

$L_{1,1}$	$W_{1,2}$	$W_{2,1}$	$W_{1,2} \vee W_{2,1}$	$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$
F	F	F	F	F
F	F	T	T	F
F	T	F	T	F
F	T	T	T	F
T	F	F	F	F
T	F	T	T	T
T	T	F	T	T
T	T	T	T	T

Models

$$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$$

$L_{1,1}$	$W_{1,2}$	$W_{2,1}$	$W_{1,2} \vee W_{2,1}$	$L_{1,1} \wedge (W_{1,2} \vee W_{2,1})$
F	F	F	F	F
F	F	T	T	F
F	T	F	T	F
F	T	T	T	F
T	F	F	F	F
T	F	T	T	T
T	T	F	T	T
T	T	T	T	T

Models

$$\{\ L_{1,1},\ (W_{1,2} \vee W_{2,1})\ \}$$

$L_{1,1}$	$W_{1,2}$	$W_{2,1}$	$W_{1,2} \vee W_{2,1}$	$L_{1,1}$, $(W_{1,2} \vee W_{2,1})$
F	F	F	F	F
F	F	T	T	F
F	T	F	T	F
F	T	T	T	F
T	F	F	F	F
T	F	T	T	T
T	T	F	T	T
T	T	T	T	T

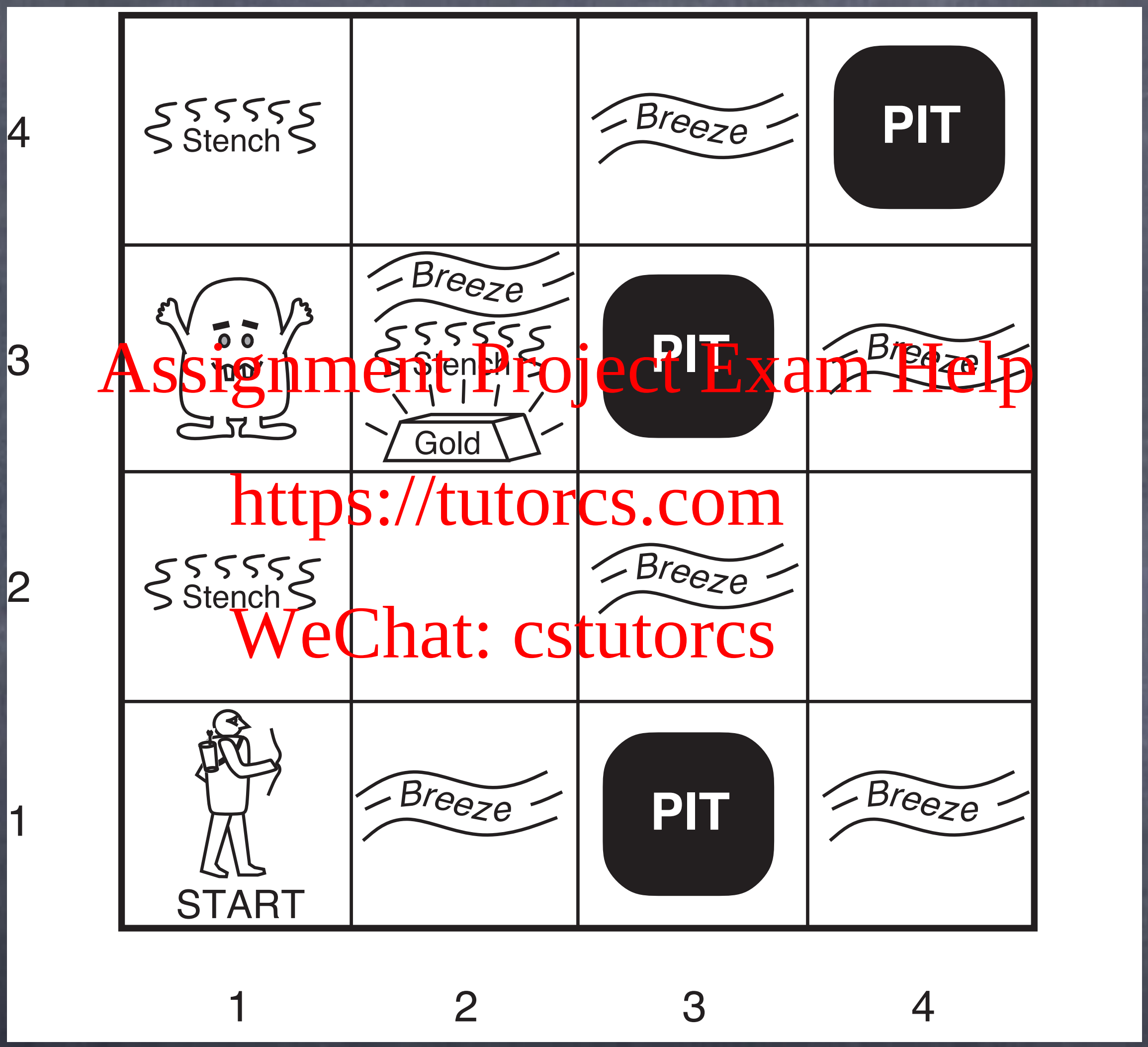
Unsatisfiable

$$L_{1,1} \wedge \neg L_{1,1}$$

$L_{1,1}$	$\neg L_{1,1}$	$L_{1,1} \wedge \neg L_{1,1}$
F	T	F
T	F	F

Propositional Logic and Boolean CSPs

- Every sentence of propositional logic represents a Boolean function of its atomic propositions
- Each assignment of true or false to the atomic propositions is one possible world
- That world is consistent with the sentence if the function denoted by the sentence is true given that assignment



4
3
2
1
Stench
Breeze
PIT
Breeze
Stench
Gold
PIT
Breeze
Stench
Breeze
START
Breeze
PIT
Breeze
1
2
3
4

$P_{1,1}$, $P_{1,2}$, … : There is a pit at [1,1], [1,2], …

$W_{1,1}$, $W_{1,2}$, … : The wumpus is at [1,1], [1,2], …

…

$B_{1,1}$, $B_{1,2}$, … : You perceived a breeze at [1,1], …

$S_{1,1}$, $S_{1,2}$, … : You perceived a stench at [1,1], …

…

$L_{1,1}$, $L_{1,2}$, … : The agent is at [1,1], …

$Facing_N$, $Facing_S$, $Facing_E$, $Facing_W$

Background Knowledge

Background Knowledge

- You perceive a breeze in a room if that room is adjacent to a room that contains a pit

Background Knowledge

- You perceive a breeze in a room only if that room is adjacent to a room that contains a pit

Background Knowledge

- You perceive a breeze in a room if and only if that room is adjacent to a room that contains a pit
- Rooms adjacent to pits have breezes.

Background Knowledge

- You perceive a breeze in a room if and only if that room is adjacent to a room that contains a pit
- Rooms adjacent to pits have breezes.

$B_{i,j} \Leftrightarrow P_{k,l}$ for some room $[k,l]$ adjacent to $[i,j]$

Background Knowledge

- You perceive a breeze in a room if and only if that room is adjacent to a room that contains a pit
- Rooms adjacent to pits have breezes.
- You perceive a breeze in [1,1] iff there is a pit in [1,2] or [2,1]

$$B_{1,1} \Leftrightarrow P_{1,2} \vee P_{2,1}$$

Background Knowledge

- You perceive a breeze in a room if and only if that room is adjacent to a room that contains a pit

- Rooms adjacent to pits have breezes.
- You perceive a breeze in [1,1] iff there is a pit in [1,2] or [2,1]
- You perceive a breeze in [1,2] iff there is a pit in [1,1] or [2,2] or [3,1]

$$B_{1,2} \Leftrightarrow P_{1,1} \vee P_{2,2} \vee P_{3,1}$$

Background Knowledge

$B_{1,1} \Leftrightarrow P_{1,2} \vee P_{2,1}$

$B_{1,2} \Leftrightarrow P_{1,1} \vee P_{2,2} \vee P_{3,1}$

$B_{2,2} \Leftrightarrow P_{1,2} \vee P_{2,3} \vee P_{3,2} \vee P_{2,1}$

...

Background Knowledge

$B_{1,1} \Leftrightarrow P_{1,2} \vee P_{2,1}$

$B_{1,2} \Leftrightarrow P_{1,1} \vee P_{2,2} \vee P_{3,1}$

$B_{2,2} \Leftrightarrow P_{1,2} \vee P_{2,3} \vee P_{3,2} \vee P_{2,1}$

...

$S_{1,1} \Leftrightarrow W_{1,2} \vee W_{2,1}$

$S_{1,2} \Leftrightarrow W_{1,1} \vee W_{2,2} \vee W_{3,1}$

$S_{2,2} \Leftrightarrow W_{1,2} \vee W_{2,3} \vee W_{3,2} \vee W_{2,1}$

...

Background Knowledge

- A room is safe ("OK") if and only if it does not contain either a pit or the wumpus

Background Knowledge

- A room is safe ("OK") if and only if it does not contain either a pit or the wumpus

$$OK_{1,1} \Leftrightarrow \neg(P_{1,1} \lor W_{1,1})$$

$$OK_{1,2} \Leftrightarrow \neg(P_{1,2} \lor W_{1,2})$$

$$OK_{2,1} \Leftrightarrow \neg(P_{2,1} \lor W_{2,1})$$

$$\ldots$$

Background Knowledge

- There is exactly one wumpus.
- The wumpus is in exactly one room.

Background Knowledge

- There is exactly one wumpus.
- The wumpus is in exactly one room.
- There is a wumpus in at least one of the rooms AND a wumpus cannot be in two rooms.

Background Knowledge

- There is exactly one wumpus.
- The wumpus is in exactly one room.
- There is a wumpus in at least one of the rooms AND a wumpus cannot be in two rooms.

$$W_{1,1} \vee W_{1,2} \vee \ldots \vee W_{3,4} \vee W_{4,4}$$

$$\neg(W_{1,1} \wedge W_{1,2}), \neg(W_{1,1} \wedge W_{1,3}), \ldots, \neg(W_{3,4} \wedge W_{4,4})$$

Background Knowledge

$B_{1,1} \Leftrightarrow P_{1,2} \vee P_{2,1}$

$B_{1,2} \Leftrightarrow P_{1,1} \vee P_{2,2} \vee P_{3,1}$

$B_{2,2} \Leftrightarrow P_{1,2} \vee P_{2,3} \vee P_{3,2} \vee P_{2,1}$

...

$S_{1,1} \Leftrightarrow W_{1,2} \vee W_{2,1}$

$S_{1,2} \Leftrightarrow W_{1,1} \vee W_{2,2} \vee W_{3,1}$

$S_{2,2} \Leftrightarrow W_{1,2} \vee W_{2,3} \vee W_{3,2} \vee W_{2,1}$

...

$OK_{1,1} \Leftrightarrow \neg(P_{1,1} \vee W_{1,1})$

$OK_{1,2} \Leftrightarrow \neg(P_{1,2} \vee W_{1,2})$

$OK_{2,1} \Leftrightarrow \neg(P_{2,1} \vee W_{2,1})$

...

$$W_{1,1} \vee W_{1,2} \vee \ldots \vee W_{3,4} \vee W_{4,4}$$

$$\neg(W_{1,1} \wedge W_{1,2}),\ \neg(W_{1,1} \wedge W_{1,3}),\ \ldots,\ \neg(W_{3,4} \wedge W_{4,4})$$

Knowledge-Based Agents

Knowledge-Based Agents

- Maintain a knowledge base (KB) of sentences believed to be true

Knowledge-Based Agents

- Maintain a knowledge base (KB) of sentences believed to be true
- Perceives the world (and updates its beliefs)

Knowledge-Based Agents

- Maintain a knowledge base (KB) of sentences believed to be true
- Perceives the world (and updates its beliefs)
- Infers what to do next

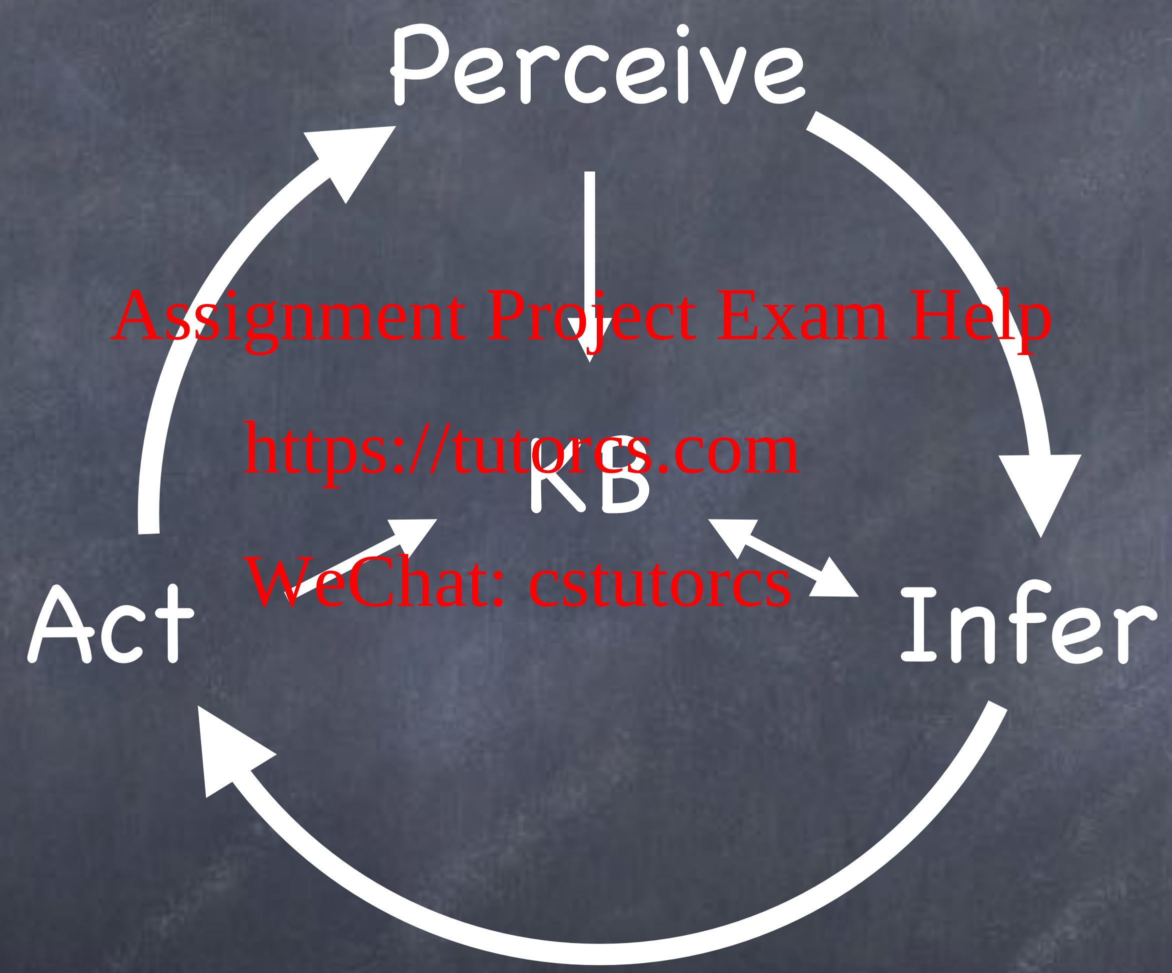
Knowledge-Based Agents

- Maintain a knowledge base (KB) of sentences believed to be true

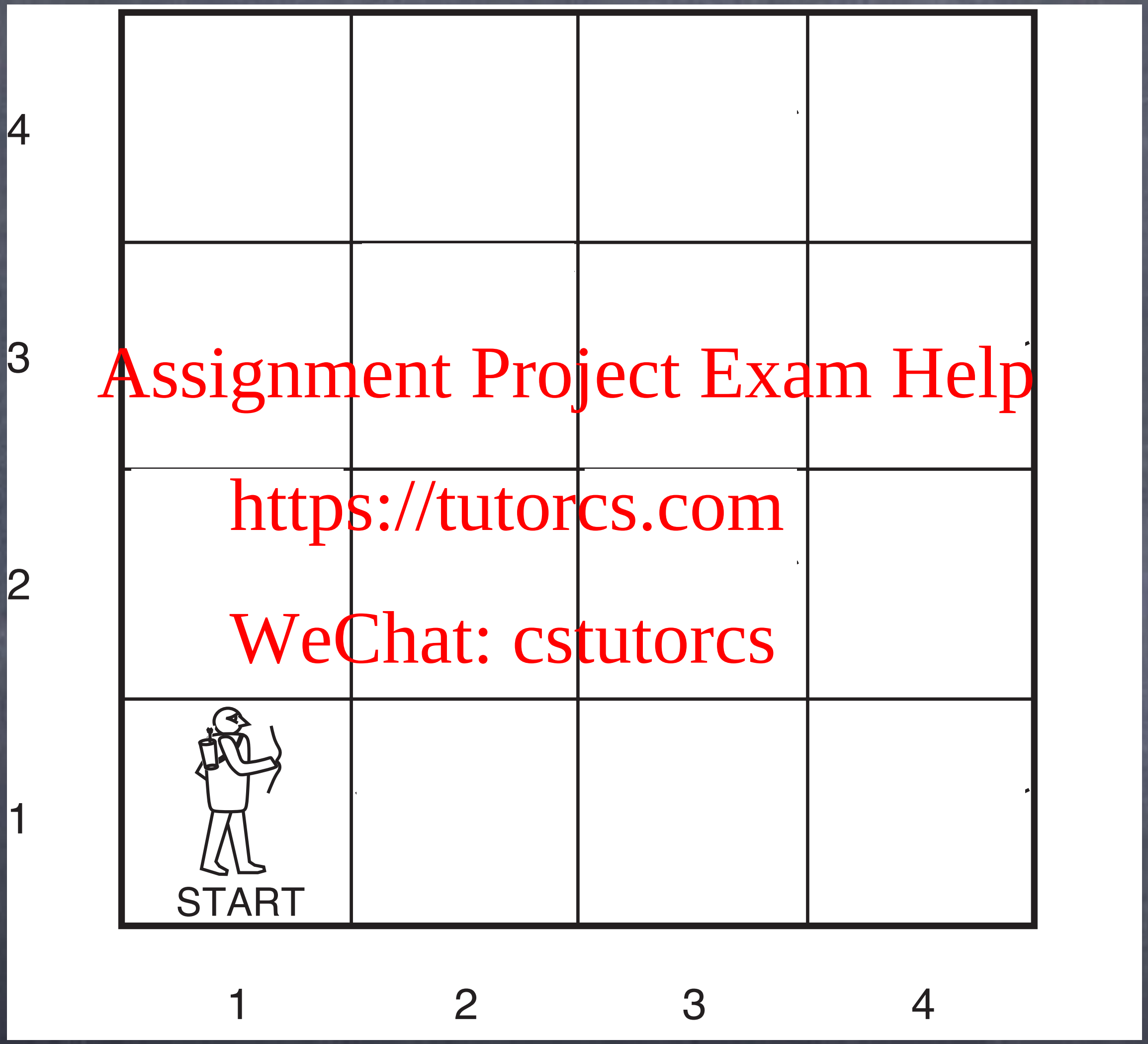
- Perceives the world (and updates its beliefs)

- Infers what to do next
- Performs an action (and updates its beliefs)

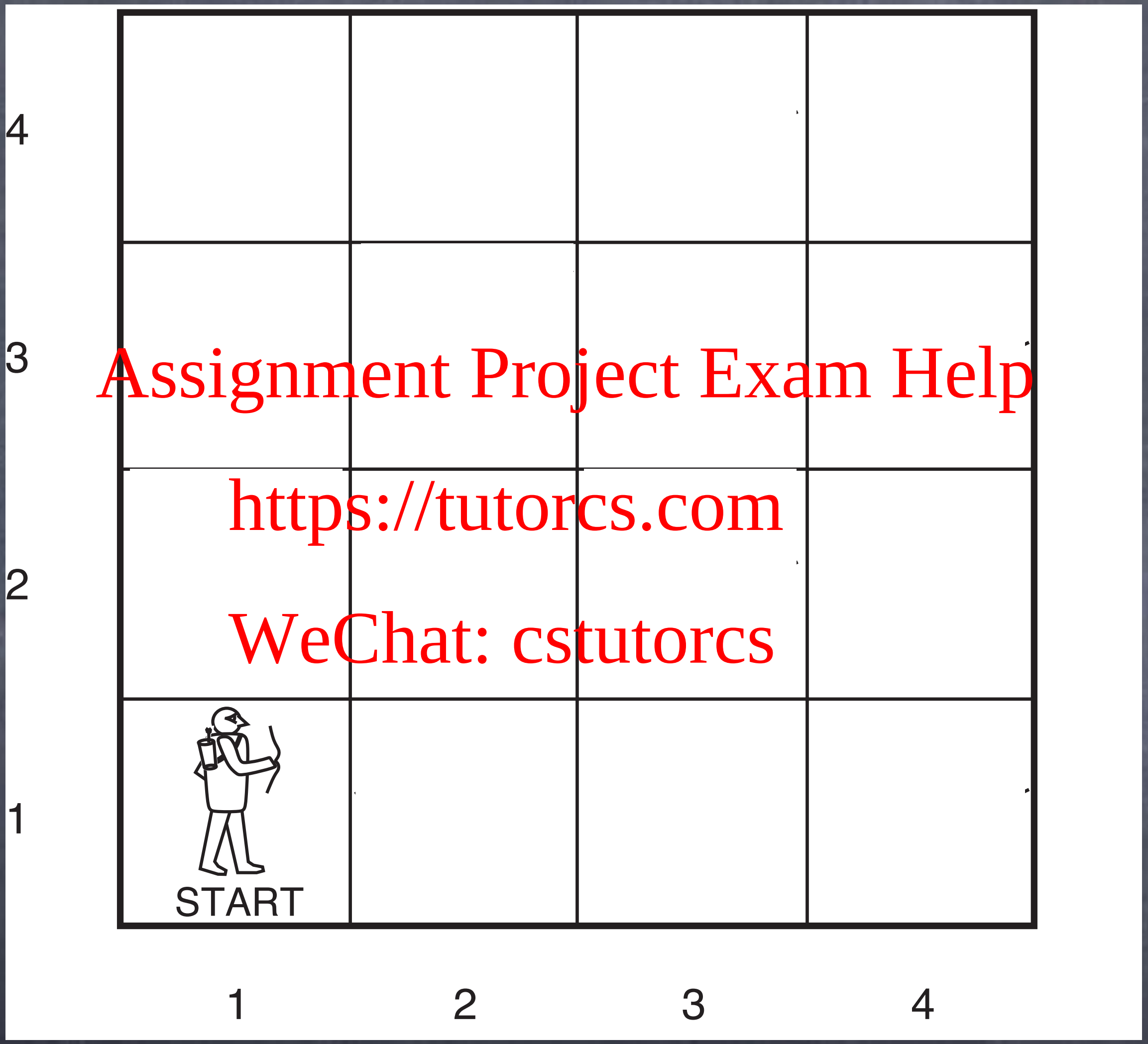
Knowledge-Based Agents



Knowledge-Based Agents



Knowledge-Based Agents

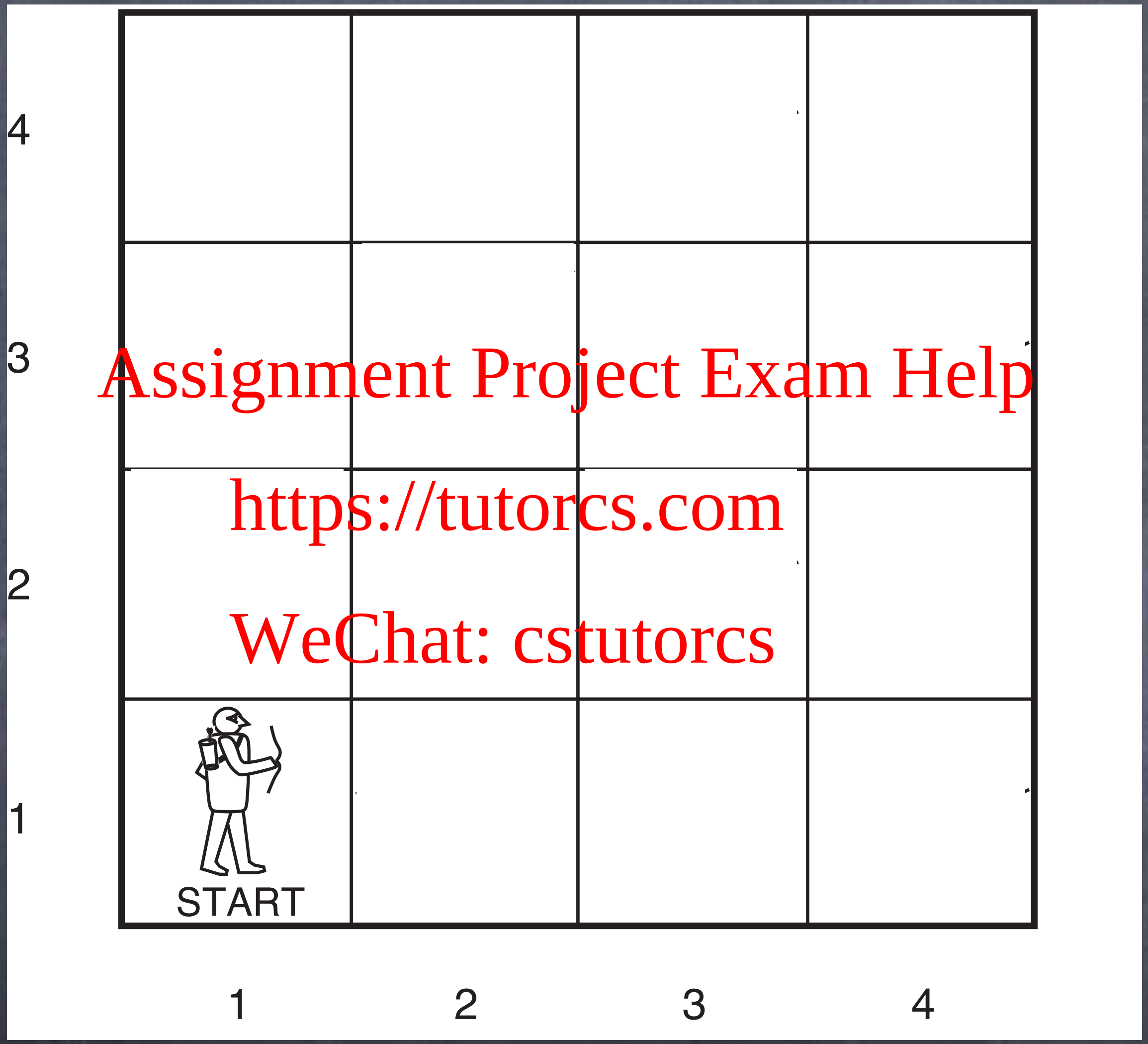


$$L_{1,1},\ \neg L_{1,2},\ \ldots\ \neg L_{3,4},\ \neg L_{4,4}$$

$$FacingE,\ \neg FacingN,\ \neg FacingS,\ \neg FacingW$$

$$OK_{1,1}$$

Perception



$$\neg B_{1,1},\ \neg S_{1,1}$$

Inference

- Given what I know…
  - What should I do?

See AIMA 7.7

Inference

- Given what I know...
  - What should I do?
  - Is room [2,1] or [1,2] safe?

Inference

- Given what I know...
  - What should I do?
  - Is room [2,1] or [1,2] safe?
  - Is room [2,1] safe?

Inference

- Given what I know...
  - What should I do?
  - Is room [2,1] or [1,2] safe?
  - Is room [2,1] safe?
  - Is there no pit in room [2,1]?

Inference

Given what I know... Is there no pit in room [2,1]?

KB $\neg P_{2,1}?$

Inference

Given what I know... Is there no pit in room [2,1]?

R_1: $\neg P_{1,1}$

R_2: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$

$\neg P_{2,1}?$

$\neg B_{1,1}$

Inference

Possible Worlds

$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$B_{1,1}$				
F	F	F	F				
F	F	F	T				
F	F	T	T				
$\cdots$							
T	T	F	T				
T	T	T	F				
T	T	T	T				

Inference

Knowledge

$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$B_{1,1}$	R_1	R_2	$\neg B_{1,1}$	$\neg P_{2,1}$
F	F	F	F	T	T	T	
F	F	F	T	T	F	F	
F	F	T	T	T	F	T	
$\ldots$				$\ldots$			
T	T	F	T	F	T	F	
T	T	T	F	F	F	T	
T	T	T	T	F	T	F	

Impossible Worlds

Knowledge

$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$B_{1,1}$	R_1	R_2	$\neg B_{1,1}$	$\neg P_{2,1}$
F	F	F	F	T	T	T	
F	F	F	T	T	F	F	
F	F	T	T	T	F	T	
...				...			
T	T	F	T	F	T	F	
T	T	T	F	F	F	T	
T	T	T	T	F	T	F	

Models

Knowledge

$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$B_{1,1}$	R_1	R_2	$\neg B_{1,1}$	$\neg P_{2,1}$
F	F	F	F	T	T	T	
~~F~~	~~F~~	~~F~~	~~T~~	~~T~~	~~F~~	~~F~~	
~~F~~	~~F~~	~~T~~	~~T~~	~~T~~	~~F~~	~~T~~	
...				...			
~~T~~	~~T~~	~~F~~	~~T~~	~~F~~	~~T~~	~~F~~	
~~T~~	~~T~~	~~T~~	~~F~~	~~F~~	~~F~~	~~T~~	
~~T~~	~~T~~	~~T~~	~~T~~	~~F~~	~~T~~	~~F~~	

Inference

Query

$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$B_{1,1}$	R_1	R_2	$\neg B_{1,1}$	$\neg P_{2,1}$
F	F	F	F	T	T	T	T
~~F~~	~~F~~	~~F~~	~~T~~	~~T~~	~~F~~	~~F~~	
~~F~~	~~F~~	~~T~~	~~T~~	~~T~~	~~F~~	~~T~~	
$\dots$				$\dots$			
~~T~~	~~T~~	~~F~~	~~T~~	~~F~~	~~T~~	~~F~~	
~~T~~	~~T~~	~~T~~	~~F~~	~~F~~	~~F~~	~~T~~	
~~T~~	~~T~~	~~T~~	~~T~~	~~F~~	~~T~~	~~F~~	

Entailment

- α entails β : $\alpha \vDash \beta$
  - Every model of α is also a model of β

Entailment

- α entails β : $\alpha \vDash \beta$
  - Every model of α is also a model of β
  - Whenever α is true, so is β
  - β is true in every world consistent with α
  - $Models(\alpha) \subseteq Models(\beta)$
  - β logically follows from α

Entailment

- Is very conservative
  - Only accepts a conclusion that is guaranteed to be true whenever the premises are true

Entailment

- Is very conservative
  - Only accepts a conclusion that is guaranteed to be true whenever the premises are true
- If β is false in every model of α, then $\alpha \models \neg\beta$

Entailment

- Is very conservative
  - Only accepts a conclusion that is guaranteed to be true whenever the premises are true
- If β is false in every model of α, then $\alpha \vDash \neg\beta$
- Otherwise: don't know!

Entailment

Given what I know... Is there no pit in room [2,1]?

R_1: $\neg P_{1,1}$

R_2: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$

$\neg P_{2,1}$?

$\neg B_{1,1}$

$$KB \models \neg P_{2,1}$$

Entailment

Given what I know... Is there no pit in room [1,2]?

R_1: $\neg P_{1,1}$

R_2: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$

R_3: $B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$

$\neg B_{1,1}$

$B_{2,1}$

$\neg P_{1,2}$?

Inference

$B_{1,1}$	$B_{1,2}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R_1	R_2	R_3	R_4	R_5	KB
F	F	F	F	F	F	F	T	T	T	T	F	F
F	F	F	F	F	F	T	T	T	F	T	F	F
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
F	T	F	F	F	F	F	T	T	F	T	T	F
F	T	F	F	F	F	T	T	T	T	T	T	T
F	T	F	F	F	T	F	T	T	T	T	T	T
F	T	F	F	F	T	T	T	T	T	T	T	T
F	T	F	F	T	F	F	T	F	F	T	T	F
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
T	T	T	T	T	T	T	F	T	T	F	T	F

Inference

Possible Worlds

$B_{1,1}$	$B_{1,2}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R_1	R_2	R_3	R_4	R_5	KB
F	F	F	F	F	F	F	T	T	T	T	F	F
F	F	F	F	F	F	T	T	T	F	T	F	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
F	T	F	F	F	F	F	T	T	F	T	T	F
F	T	F	F	F	F	T	T	T	T	T	T	T
F	T	F	F	F	T	F	T	T	T	T	T	T
F	T	F	F	F	T	T	T	T	T	T	T	T
F	T	F	F	T	F	F	T	F	F	T	T	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
T	T	T	T	T	T	T	F	T	T	F	T	F

Inference

Knowledge

$B_{1,1}$	$B_{1,2}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R_1	R_2	R_3	R_4	R_5	KB
F	F	F	F	F	F	F	T	T	T	T	F	F
F	F	F	F	F	F	T	T	T	F	T	F	F
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
F	T	F	F	F	F	F	T	T	F	T	T	F
F	T	F	F	F	F	T	T	T	T	T	T	T
F	T	F	F	F	T	F	T	T	T	T	T	T
F	T	F	F	F	T	T	T	T	T	T	T	T
F	T	F	F	T	F	F	T	F	F	T	T	F
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
T	T	T	T	T	T	T	F	T	T	F	T	F

Impossible Worlds

$B_{1,1}$	$B_{1,2}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R_1	R_2	R_3	R_4	R_5	KB
F	F	F	F	F	F	F	T	T	T	T	F	F
F	F	F	F	F	F	T	T	T	F	T	F	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
F	T	F	F	F	F	F	T	T	F	T	T	F
F	T	F	F	F	F	T	T	T	T	T	T	T
F	T	F	F	F	T	F	T	T	T	T	T	T
F	T	F	F	F	T	T	T	T	T	T	T	T
F	T	F	F	T	F	F	T	F	F	T	T	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
T	T	T	T	T	T	T	F	T	T	F	T	F

Models

$B_{1,1}$	$B_{1,2}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R_1	R_2	R_3	R_4	R_5	KB
F	F	F	F	F	F	F	T	T	T	T	F	F
F	F	F	F	F	F	T	T	T	F	T	F	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
F	T	F	F	F	F	F	T	T	F	T	T	F
F	T	F	F	F	F	T	T	T	T	T	T	T
F	T	F	F	F	T	F	T	T	T	T	T	T
F	T	F	F	F	T	T	T	T	T	T	T	T
F	T	F	F	T	F	F	T	F	F	T	T	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
T	T	T	T	T	T	T	F	T	T	F	T	F

Inference

$KB \models \neg P_{1,2}$

$B_{1,1}$	$B_{1,2}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R_1	R_2	R_3	R_4	R_5	KB
F	F	F	F	F	F	F	T	T	T	T	F	F
F	F	F	F	F	F	T	T	T	F	T	F	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
F	T	F	F	F	F	F	T	T	F	T	T	F
F	T	F	$\boxed{F}$	F	F	T	T	T	T	T	T	T
F	T	F	$\boxed{F}$	F	T	F	T	T	T	T	T	T
F	T	F	$\boxed{F}$	F	T	T	T	T	T	T	T	T
F	T	F	F	T	F	F	T	F	F	T	T	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
T	T	T	T	T	T	T	F	T	T	F	T	F

Inference

$KB \not\models P_{2,2}$

$KB \not\models \neg P_{2,2}$

$B_{1,1}$	$B_{1,2}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R_1	R_2	R_3	R_4	R_5	KB
F	F	F	F	F	F	F	T	T	T	T	F	F
F	F	F	F	F	F	T	T	T	F	T	F	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
F	T	F	F	F	F	F	T	T	F	T	T	F
F	T	F	F	F	$\boxed{F}$	T	T	T	T	T	T	T
F	T	F	F	F	$\boxed{T}$	F	T	T	T	T	T	T
F	T	F	F	F	$\boxed{T}$	T	T	T	T	T	T	T
F	T	F	F	T	F	F	T	F	F	T	T	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
T	T	T	T	T	T	T	F	T	T	F	T	F

Entailment

Given what I know... Is there no pit in room [1,2]?

R_1: $\neg P_{1,1}$

R_2: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$

$\neg P_{1,2}?$

R_3: $B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$

$\neg B_{1,1}$

$B_{2,1}$

$KB \models \neg P_{1,2}$

$KB \not\models P_{2,1}$

$KB \not\models \neg P_{2,1}$

Inference

- Given what I know...
  - What should I do?
  - Is room [2,1] or [2,1] safe?
  - Is room [2,1] safe?

Inference

Given what I know...

Is room [2,1] safe?

R_1: $OK_{1,1}$

R_2: $OK_{1,1} \Leftrightarrow \neg(P_{1,1} \vee W_{1,1})$

$OK_{2,1}$?

R_3: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$

R_4: $S_{1,1} \Leftrightarrow (W_{1,2} \vee W_{2,1})$

R_5: $OK_{2,1} \Leftrightarrow \neg(P_{2,1} \vee W_{2,1})$

$\neg B_{1,1}$, $\neg S_{1,1}$

Inference

Possible Worlds

$P_{1,1}$	$P_{1,2}$	$\cdots$	$OK_{1,1}$	$OK_{2,1}$	R_1	R_2	$\cdots$	$\neg B_{1,1}$	$\neg S_{1,1}$	$OK_{2,1}$
F	F	...	F	F						
...										
F	F	...	T	T						
...										
T	T	...	F	T						
T	T	...	T	F						
T	T	...	T	T						

Inference

Knowledge

$P_{1,1}$	$P_{1,2}$	$\cdots$	$OK_{1,1}$	$OK_{2,1}$	R_1	R_2	$\cdots$	$\neg B_{1,1}$	$\neg S_{1,1}$	$OK_{2,1}$
F	F	…	F	F	F	…	…	…	T	
…					…					
F	F	…	T	T	T	T	…	…	T	
…					…					
T	T	…	F	T	F	…	…	…	F	
T	T	…	T	F	T	F	…	…	T	
T	T	…	T	T	T	T	…	…	F	

Impossible Worlds

$P_{1,1}$	$P_{1,2}$	$\cdots$	$OK_{1,1}$	$OK_{2,1}$	R_1	R_2	$\cdots$	$\neg B_{1,1}$	$\neg S_{1,1}$	$OK_{2,1}$
~~F~~	~~F~~	~~...~~	~~F~~	~~F~~	~~F~~		~~...~~	~~...~~	~~T~~	
~~...~~										
F	F	...	T	T	T	T	...	...	T	
~~...~~					~~...~~					
~~T~~	~~T~~	~~...~~	~~F~~	~~T~~	~~F~~	~~...~~	~~...~~	~~...~~	~~F~~	
~~T~~	~~T~~	~~...~~	~~T~~	~~F~~	~~T~~	~~F~~	~~...~~	~~...~~	~~T~~	
~~T~~	~~T~~	~~...~~	~~T~~	~~T~~	~~T~~	~~T~~	~~...~~	~~...~~	~~F~~	

Inference

Query

$P_{1,1}$	$P_{1,2}$	$\cdots$	$OK_{1,1}$	$OK_{2,1}$	R_1	R_2	$\cdots$	$\neg B_{1,1}$	$\neg S_{1,1}$	$OK_{2,1}$
F	F	…	F	F	F	…	…	…	T	
…					…					
F	F	…	T	T	T	T	…	…	T	T
…					…					
T	T	…	F	T	F	…	…	…	F	
T	T	…	T	F	T	F	…	…	T	
T	T	…	T	T	T	T	…	…	F	

Entailment

Given what I know... Is room [2,1] safe?

R_1: $OK_{1,1}$

R_2: $OK_{1,1} \Leftrightarrow \neg(P_{1,1} \vee W_{1,1})$

$OK_{2,1}$?

R_3: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$

R_4: $S_{1,1} \Leftrightarrow (W_{1,2} \vee W_{2,1})$

R_5: $OK_{2,1} \Leftrightarrow \neg(P_{2,1} \vee W_{2,1})$

$KB \models \mathrm{OK}_{2,1}$

$\neg B_{1,1}$, $\neg S_{1,1}$

Entailment

- α entails β : $\alpha \models \beta$
  - Every model of α is also a model of β
  - Whenever α is true, so is β
  - β is true in every world consistent with α
  - $Models(\alpha) \subseteq Models(\beta)$
  - β logically follows from α

Computing Entailment

- Given knowledge α and query β
  - For every possible world w
    - If α is satisfied by w
      - If β is not satisfied by w
        - Conclude that $\alpha \not\models \beta$
  - Conclude that $\alpha \models \beta$

Model Checking

- Given knowledge α and query β
  - For every possible world w
    - If α is satisfied by w
      - If β is not satisfied by w
        - Conclude that $\alpha \not\models \beta$
  - Conclude that $\alpha \models \beta$

Computing Entailment

```
boolean TT_Entails?(KB, α)
   symbols ← proposition symbols used in KB and α
   return TT_Check_All(KB, α, symbols, {})

boolean TT_Check_All(KB, α, symbols, model)
   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})
               && TT_Check_All(KB, α, rest, model ∪ {P=false})
```

But...

Model Checking

- Given knowledge α and query β
  - For every possible world w
    - If α is satisfied by w
      - If β is not satisfied by w
        - Conclude that $\alpha \not\models \beta$
  - Conclude that $\alpha \models \beta$

Model Checking

n propositions m sentences, $O(k)$ connectives

$P_{1,1}$	$P_{1,2}$	$\cdots$	$OK_{1,1}$	$OK_{2,1}$	R_1	R_2	$\cdots$	$\neg B_{1,1}$	$\neg S_{1,1}$	$OK_{2,1}$
F	F	...	F	F	F	...	...	...	T	
...					...					
F	F	...	T	T	T	T	...	...	T	T
...					...					
T	T	...	F	T	F	...	...	...	F	
T	T	...	T	F	T	F	...	...	T	
T	T	...	T	T	T	T	...	...	F	

Model Checking

n propositions m sentences, $O(k)$ connectives

$P_{1,1}$	$P_{1,2}$	$\cdots$	$OK_{1,1}$	$OK_{2,1}$	R_1	R_2	$\cdots$	$\neg B_{1,1}$	$\neg S_{1,1}$	$OK_{2,1}$
F	F	…	F	F	F	…	…	…	T	
…					…					
F	F	…	T	T	T	T	…	…	T	T
…					…					
T	T	…	F	T	F	…	…	…	F	
T	T	…	T	F	T	F	…	…	T	
T	T	…	T	T	T	T	…	…	F	

$O(k)$

Model Checking

n propositions $\quad m$ sentences, $O(k)$ connectives

$P_{1,1}$	$P_{1,2}$	$\cdots$	$OK_{1,1}$	$OK_{2,1}$	R_1	R_2	$\cdots$	$\neg B_{1,1}$	$\neg S_{1,1}$	$OK_{2,1}$
F	F	$\cdots$	F	F	F	$\cdots$	$\cdots$	$\cdots$	T	
$\cdots$					$\cdots$					
F	F	$\cdots$	T	T	T	T	$\cdots$	$\cdots$	T	T
$\cdots$					$\cdots$					
T	T	$\cdots$	F	T	F	$\cdots$	$\cdots$	$\cdots$	F	
T	T	$\cdots$	T	F	T	F	$\cdots$	$\cdots$	T	
T	T	$\cdots$	T	T	T	T	$\cdots$	$\cdots$	F	

$O(mk)$

Model Checking

n propositions m sentences, $O(k)$ connectives

$P_{1,1}$	$P_{1,2}$	$\cdots$	$OK_{1,1}$	$OK_{2,1}$	R_1	R_2	$\cdots$	$\neg B_{1,1}$	$\neg S_{1,1}$	$OK_{2,1}$
F	F	…	F	F	F	…	…	…	T	
…					…					
F	F	…	T	T	T	T	…	…	T	T
…					…					
T	T	…	F	T	F	…	…	…	F	
T	T	…	T	F	T	F	…	…	T	
T	T	…	T	T	T	T	…	…	F	

2^n

Model Checking

n propositions $\quad m$ sentences, $O(k)$ connectives

$P_{1,1}$	$P_{1,2}$	$\cdots$	$OK_{1,1}$	$OK_{2,1}$	R_1	R_2	$\cdots$	$\neg B_{1,1}$	$\neg S_{1,1}$	$OK_{2,1}$
F	F	...	F	F	F	...	...	...	T	
...					...					
F	F	...	T	T	T	T	...	...	T	T
...					...					
T	T	...	F	T	F	...	...	...	F	
T	T	...	T	F	T	F	...	...	T	
T	T	...	T	T	T	T	...	...	F	

$O(2^n mk)$ Intractable!

Propositional Logic

- Programming language for knowledge
- Factored representation (Boolean CSP)
  - Propositions, connectives, sentences
- Possible worlds, satisfiability, models
- Entailment: $\alpha \vDash \beta$
  - Every model of α is a model of β
  - Intractable!

For next time:

AIMA Ch. 7.5