DAI-Labor
TU Berlin

# Grundlagen der Künstlichen Intelligenz

## Logikbasierte Agenten

18.11.2005

Brijnesh J Jain
bjj@dai-labor.de

A|O|T
Agententechnologien in betrieblichen Anwendungen und der Telekommunikation

# Gliederung

- ⇨ **Einleitung**
- ⇨ **Die Wumpus Welt**
- ⇨ **Aussagenlogik**
- ⇨ **Inferenz**
- ⇨ **Zusammenfassung**

## Einleitung

**Denken Menschen logisch?**

- **Selektionsaufgabe nach Watson (1966):**
  - Gegeben sind 4 Karten, von denen jede auf der einen Seite mit einem Buchstaben, auf der anderen Seite mit einer Zahl beschriftet ist.
- **Beispiel:**

  | E | K | 4 | 7 |
  |---|---|---|---|

- **Regel:**
  - Wenn auf der einen Seite der Karte ein Vokal steht, dann steht auf der anderen Seite eine gerade Zahl.
- **Aufgabe:**
  - Welche Karten müssen für die Überprüfung der Regel umgedreht werden?

# Einleitung

**Denken Menschen logisch?**

- **Antwort:**
  - Menschen, so scheint es, wissen Dinge über die Umwelt und können daraus neues Wissen folgern und „intelligente“ Aktionen ableiten
- **Ziel dieser VL:**
  - Agenten, die Wissen repräsentieren und neues Wissen schlussfolgern können
- **Zentrale Konzepte:**
  - Maschinengerechte Repräsentation von Wissen
  - Formale Mechanismen für Schlussfolgern (*reasoning*)

# Einleitung

## Realisierung der Konzepte

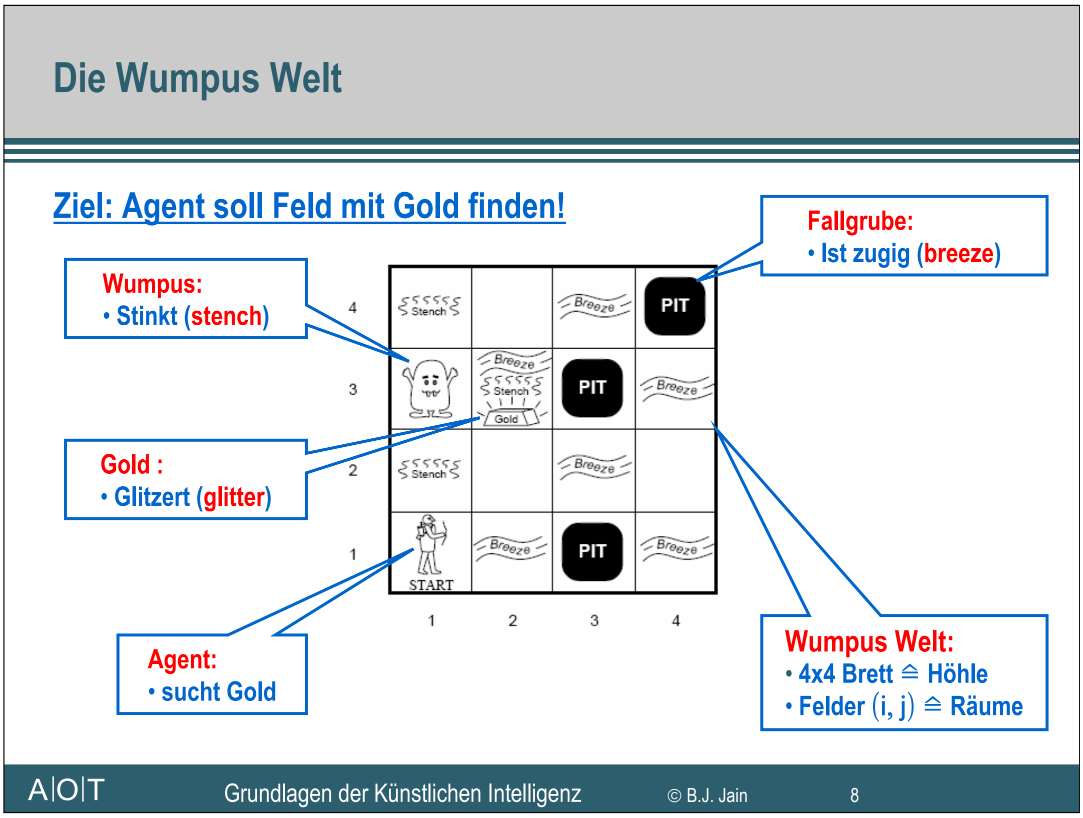
- **Logikbasierte Agenten**
  - Agenten, deren Wissensbasis und Mechanismen zum Schlussfolgern auf Formalismen der Logik beruhen
- **Logik** ist eine formale Sprache, um **Wissen** zu repräsentieren, mit dem Ziel bestimmte **Schlussfolgerungen** ziehen zu können.
- **Konzepte dieser Vorlesung**
  - Aussagenlogik als formale Sprache für die Wissensrepräsentation
  - Resolutionsmethode als Mechanismus zum logischen Schlussfolgern

## Gliederung

- ⇨ Einleitung
- ⇨ **Die Wumpus Welt**
- ⇨ Aussagenlogik
- ⇨ Inferenz
- ⇨ Zusammenfassung

## Die Wumpus Welt

- Entwickelt von Gregory Yob (1975) als Computerspiel
- Später adaptiert von Michael Genesereth als Testbettumgebung für logikbasierte Agenten
- Diskutiert in Russell & Norvig als motivierendes Beispiel für Konzepte der Aussagenlogik und Inferenzmechanismen
- Hier: Abgespeckte Version, um Konzepten der Aussagenlogik und Inferenzmechanismen zu illustrieren

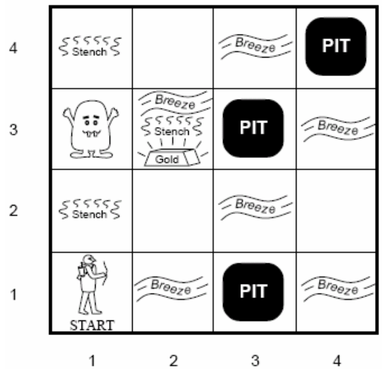



Die Funktionsweise eines logikbasierten Agenten kann sehr gut mit Hilfe der Wumpus-Welt nachvollzogen werden.

Die Wumpus-Welt ist einem frühen Computerspiel entliehen: Der Agent soll ein Höhlensystem erkunden und das dort versteckte Gold finden. In dem Höhlensystem versteckt sich ein Monster, der Wumpus, dem der Agent besser nicht begegnen sollte. Das Höhlensystem besteht aus Räumen, die über Gänge miteinander verbunden sind. Neben dem Wumpus-Raum gibt es Räume ohne besondere Merkmale und Räume mit Fallgruben.

# Wumpus Welt

- ⇨ **Umgebung:**
  - 4x4 Gitter
  - Nachbarfelder eines Feldes
    - oben, unten, links, rechts
  - Agent und Wumpus im gleichen Feld
    - Wumpus verspeist Agenten
  - Agent und Fallgrube im gleichen Feld
    - Agent fällt in Fallgrube
  - Nachbarfelder von Wumpus stinken
  - Nachbarfelder von Grube winden
  - Feld von Gold glitzert



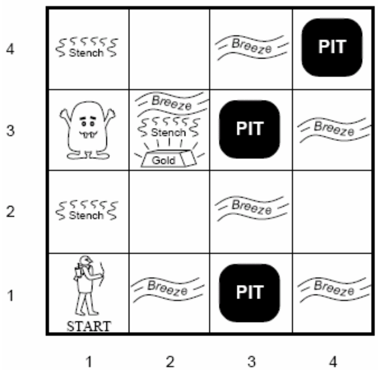
# Wumpus Welt

⇨ **Wahrnehmung des Agenten:**

- Wind (‚B‘, breeze)
- Gestank (‚S‘, stench)
- Glitzern (‚G‘, glitter)

⇨ **Aktionen:**

- Gehe zum Nachbarfeld
  - d.h. oben, unten, links, rechts
- **Bem.:**
  - Aktionen, die aus dem Brett herausführen würden sind wirkungslos



## Die Wumpus Welt

⇨**Ziel:**

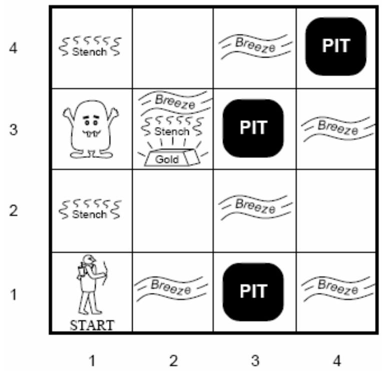
- Agent soll das Feld mit Gold in jeder denkbaren Wumpus Welt finden.
- **Bem.:** Es gibt unlösbare Konfigurationen

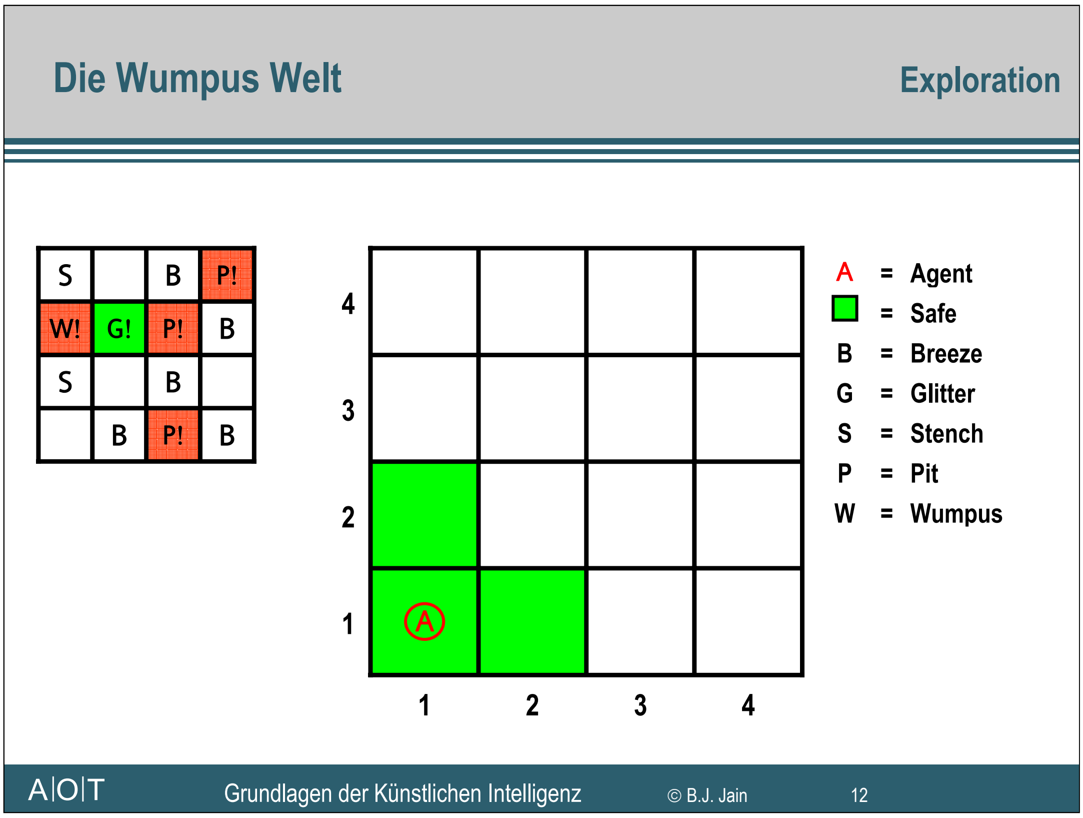
⇨**Gegeben:**

- Agent hat Wissensbasis (WB)
  - Regeln der Wumpus Welt
  - Wahrnehmung

⇨**Frage:**

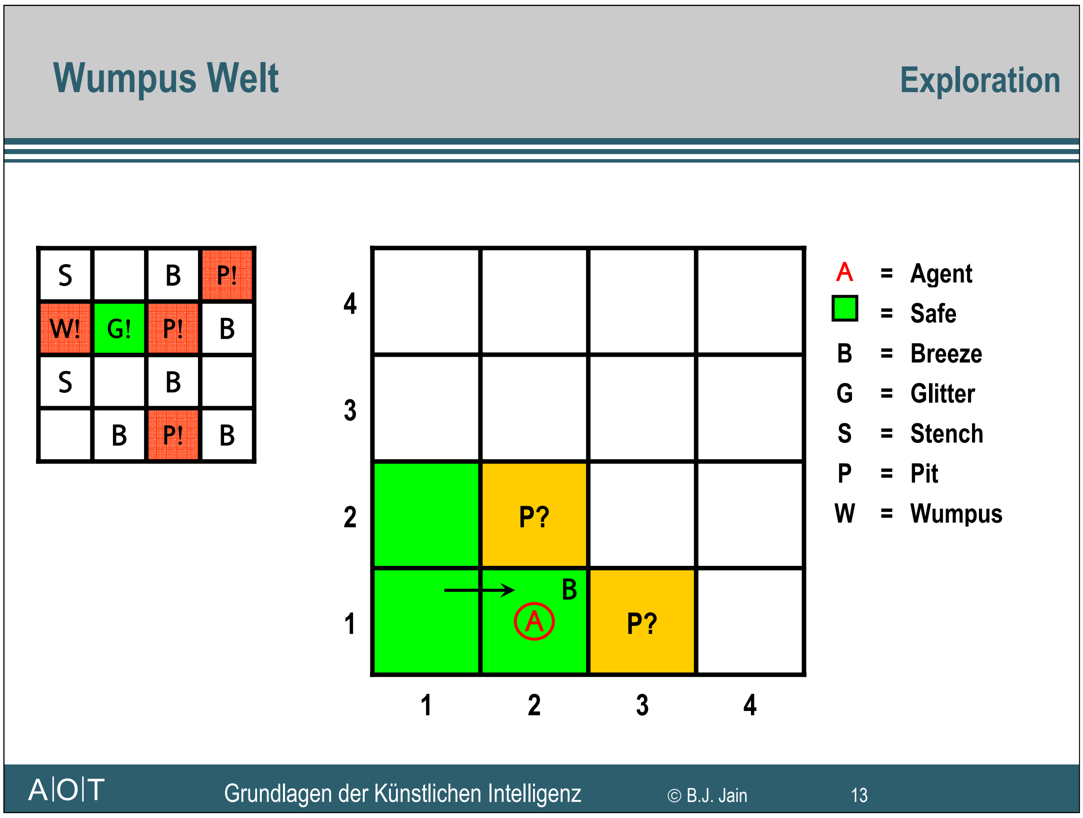
- Wie soll ein Agent in der Wumpus Welt folgern und handeln?






Zu Beginn befindet sich der Agent im Eingangsraum (1,1) und hat dort keine Wahrnehmungen. Daraus kann er schließen, dass er die benachbarten Räume (1,2) und (2,1) sicher betreten kann. Darüber hinaus kann er aus seinem Nicht-Tod schließen, dass auch sein momentaner Raum (1,1) sicher ist.

Räume, die der Agent sicher betreten kann sind grün eingefäbrt.

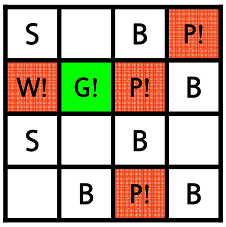


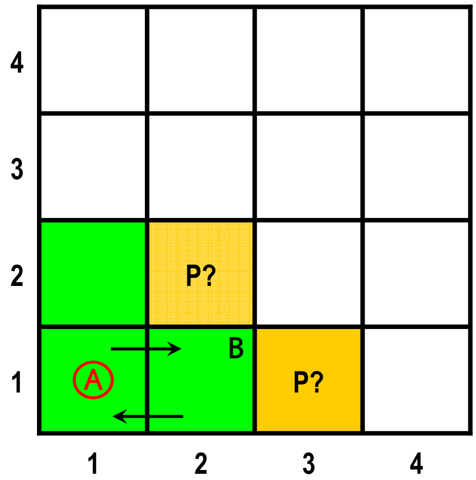

Was nun? Gegenwärtig sind alle Nachbarräume sicher und ein intelligenter Agent würde einen Raum betreten, der als sicher bekannt ist. Nehmen wir also an, der Agent entscheidet sich, nach rechts zu gehen. Dort nimmt er einen Luftzug wahr, also ist in einem (oder beiden) benachbarten und unbekannten Räumen (1,3) oder (2,2) eine Fallgrube.

Räume, die für den Agenten als unsicher gelten werden gelb eingefärbt. Mit P? deuten wir an, dass in den entsprechenden Räumen eine Fallgrube sein könnte.

## Wumpus Welt

Exploration

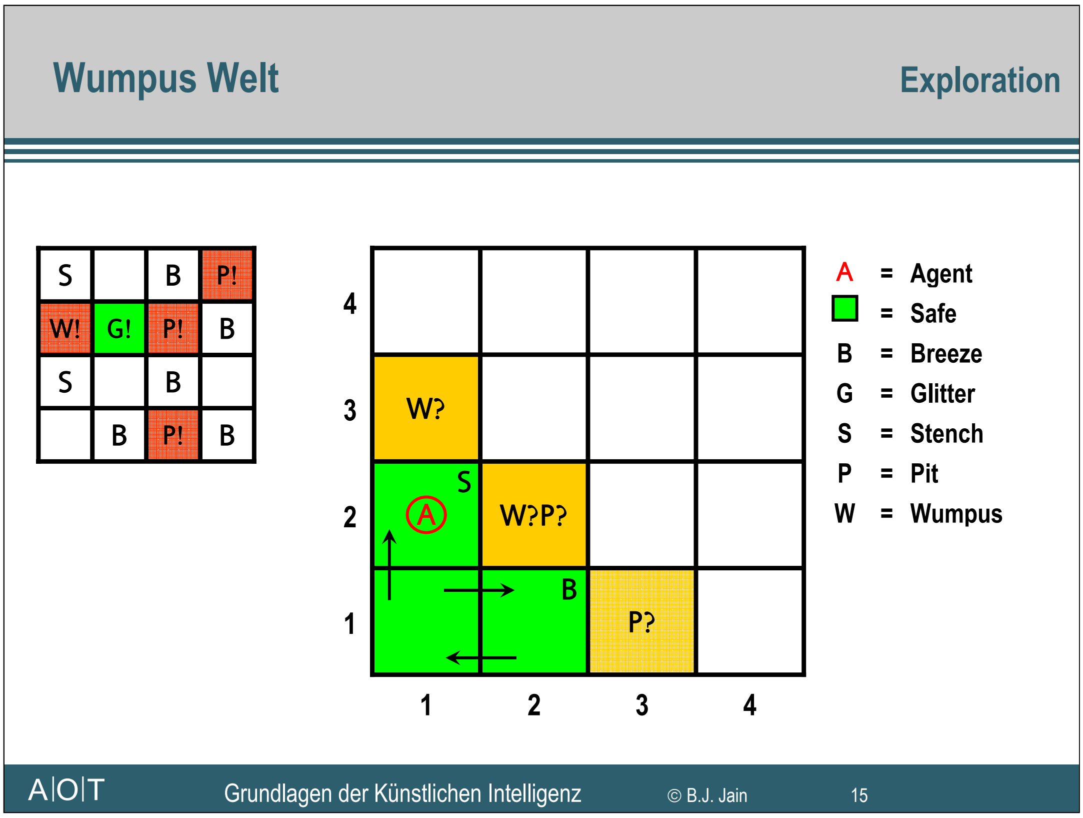


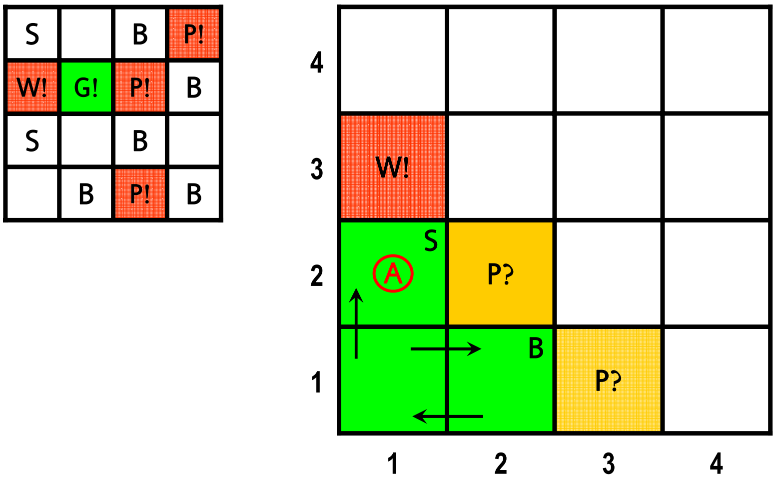

A = Agent
■ = Safe
B = Breeze
G = Glitter
S = Stench
P = Pit
W = Wumpus



Wieso sollte der Agent einen möglichen Fallgruben-Raum betreten, wenn es noch andere Möglichkeiten gibt? Der Agent wird also umkehren und den sicheren, aber unerforschten Raum (2,1) aufsuchen.



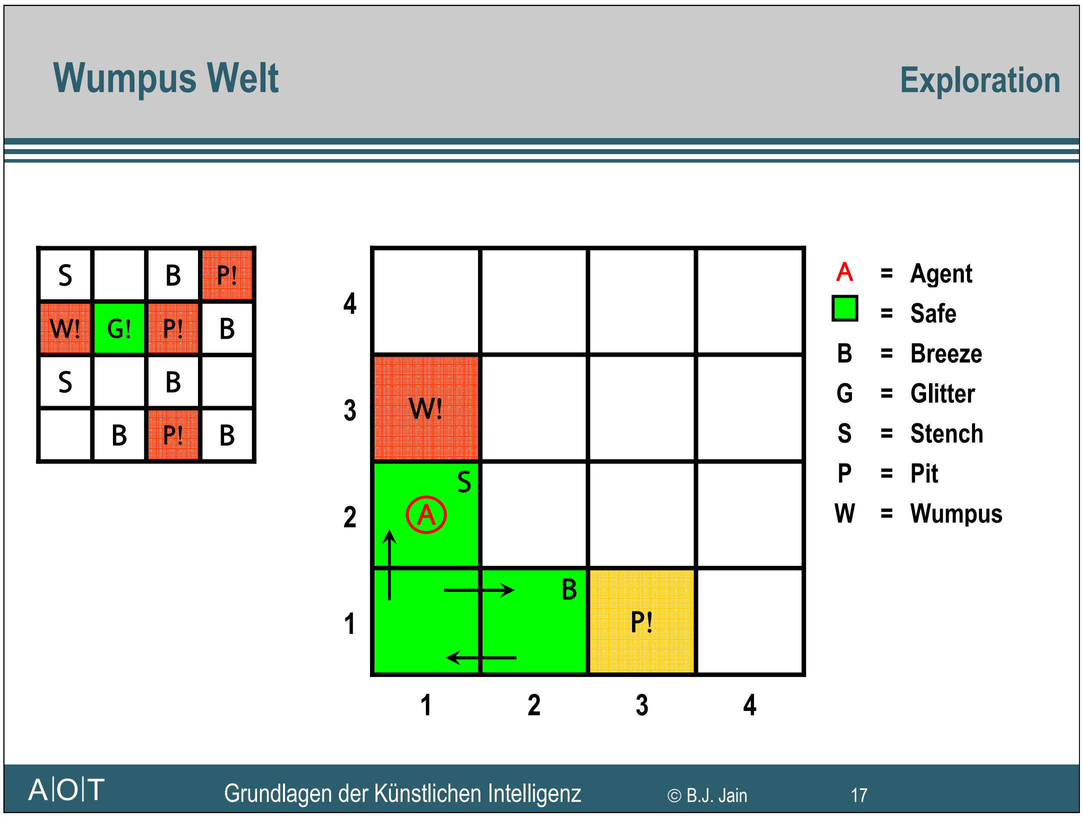

Im Raum (2,1) kann der Agent Gestank wahrnehmen. Daraus lässt sich folgern, dass in Raum (3,1) oder (2,2) der Wumpus sein könnte. Weil der Agent im Raum (1,2) keinen Gestank wahrgenommen hat, kann der Wumpus nicht im Raum (2,2) sein und muss daher im Raum (3,2) sein.



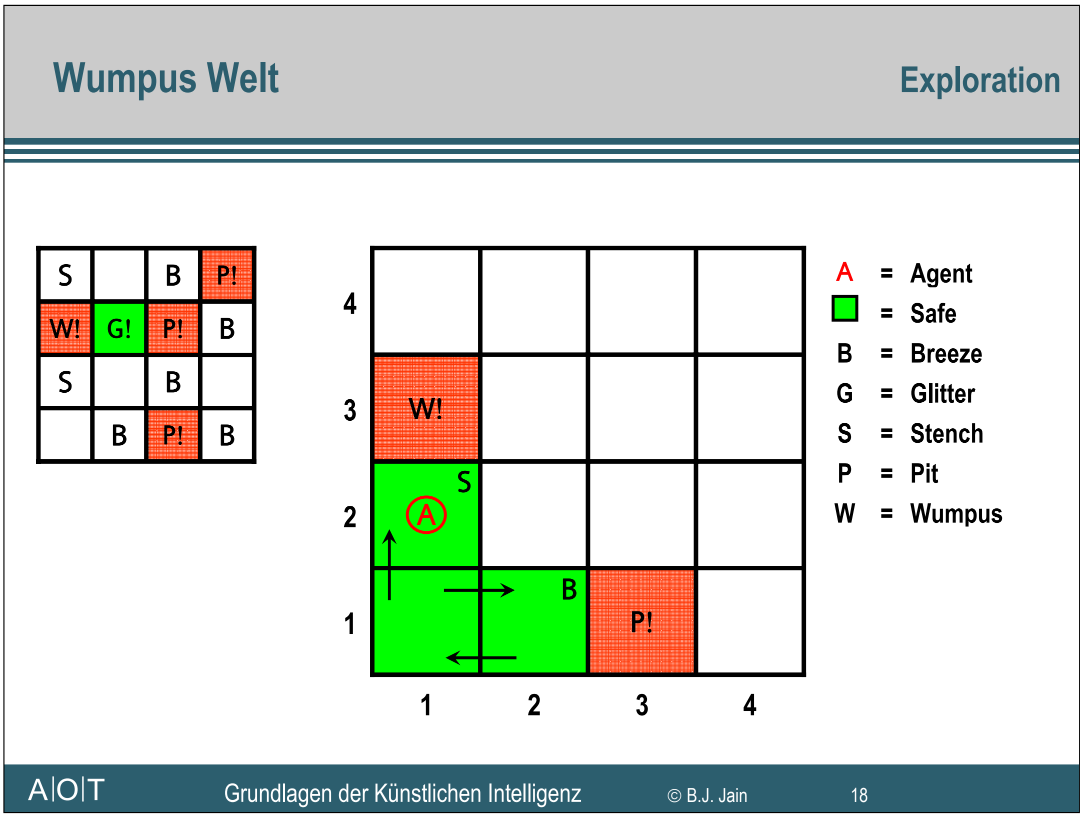

A = Agent
■ = Safe
B = Breeze
G = Glitter
S = Stench
P = Pit
W = Wumpus



Im Raum (2,1) kann der Agent keinen Luftzug wahrnehmen. Daraus lässt sich folgern, dass der Raum (2,2) keine Fallgrube haben kann.



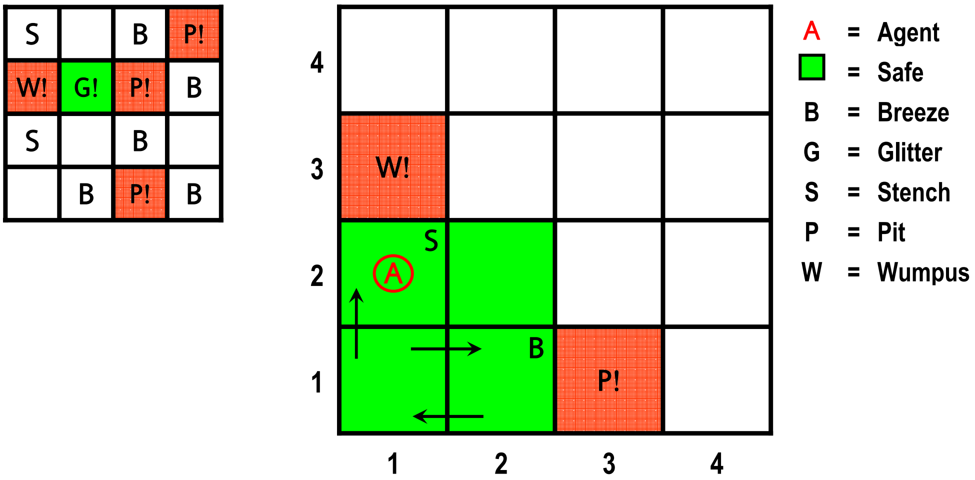

Andererseits vermutete der Agent wegen des Luftzugs in (1,2) eine Fallgrube in (2,2) oder/und (1,3). Weil in (2,2) keine Fallgrube sein kann, muss die Fallgrube also in (1,3) sein!




Weil in (2,2) weder ein Wumpus noch eine Fallgrube sein kann, ist der Raum (2,2) sicher.

## Wumpus Welt

Exploration




A|O|T Grundlagen der Künstlichen Intelligenz © B.J. Jain 19

Diese Prozedur können wir nun weiterführen bis der Agent das Gold gefunden hat.

# Die Wumpus Welt

**Was und die Wumpus Welt zeigt**

- **Intelligente Agenten benötigen**
  - Wissensbasis (WB)
  - Inferenzmechansimus

  um rationale Entscheidungen zu treffen
- **Wissensbasis**
  - Regeln der Wumpus Welt
  - Wahrnehmungen
- **Inferenzmechanismus**
  - Logisches Schließen: Wann immer der Agent aus der WB eine Schlussfolgerung zieht, so ist die Schlussfolgerung korrekt, wenn die WB korrekt ist
- **Logikbasierte Agenten = Wissensbasis + Inferenzmechanismus**

# Die Wumpus Welt

- **Problem:**
  - Wie können wir
    - die Wissensbasis
    - den Inferenzmechanismus

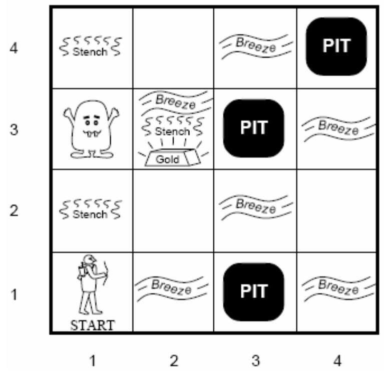
    von Agenten (maschinengerecht) formalisieren?
- **Idee:**
  - Mit Formalismus der Logik
    - Aussagenlogik zur Repräsentation von Wissen
    - Resolutionsmethode als Inferenzmechanismus

## Gliederung

## Aussagenlogik

- Aussagenlogik befasst sich mit der ...
  - logischen Bewertung von Aussagen.
  - Verknüpfung von Aussagen
- **Beispiel:**
  - Aussage 1 ist wahr:
    - Wumpus ist in Feld (3,1)
  - Aussage 2 ist falsch:
    - Gold glitzert in Feld (3,3) [false]
  - Verknüpfung ist falsch:
    - Wumpus ist in Feld (3,1)

      und
    - Gold glitzert in Feld (3,3)



## Aussagenlogik

- **Komponenten**
  - Syntax
    - definiert die wohlgeformten (zulässigen) Formeln der Sprache
    - bedeutungsleere Zeichenfolge
  - Semantik
    - definiert Bedeutung wohlgeformter Formeln in einer gegebenen „Welt“
    - Bedeutung einer Formel ist ihr Wahrheitsgehalt (true, false)
    - Eine „Welt“ kann als reale Umgebung des Agenten aufgefasst werden

- **Vokabular**
  - **Konstantensymbole:** $\top$ (wahr) und $\bot$ (falsch)
  - **Aussagensymbole:** p, q, r, s, … $\in \mathcal{A}$
  - **Junktoren:**

| Junktor | Name | Intuitive Bedeutung |
|---|---|---|
| $\neg$ | Negation | *„nicht“* |
| $\wedge$ | Konjunktion | *„und“* |
| $\vee$ | Disjunktion | *„oder“* |
| $\Rightarrow$ | Implikation | *„wenn – dann“* |
| $\Leftrightarrow$ | Äquivalenz | *„genau dann, wenn“* |

- **Aussagenlogische Formeln (wohlgeformt)**
  1. Konstanten $\top$ und $\bot$ sind Formeln
  2. Aussagesymbole sind Formeln
  3. Mit $\phi$ und $\psi$ sind folgende Ausdrücke auch Formeln:
     - $\neg \phi$
     - $\phi \wedge \psi$
     - $\phi \vee \psi$
     - $\phi \Rightarrow \psi$
     - $\phi \Leftrightarrow \psi$
  4. Alle Formeln werden nach 1.-3. gebildet.
- **Präzedenzen:**
  - Präzedenzen (hier) in Reihenfolge $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$
  - Zusätzliche Klammerungen machen Formeln trotzdem lesbarer

- **Semantik**
  - Legt fest wie man einer Formel $\phi$ eine Bedeutung zuweisen kann
  - Bedeutung von $\phi$ ist ein Wahrheitswert $1$ $(\texttt{true})$ oder $0$ $(\texttt{false})$
- **Semantik von ´Symbolen**
  - Interpretation I: $\mathcal{A} \cup \{\top, \bot\} \rightarrow \{0, 1\}$
  - Weist jedem Aussagesymbol aus $\mathcal{A}$ einen Wahrheitswert zu
  - Weist $\top$ den Wahrheitswert $1$ und $\bot$ den Wert $0$ zu
- **Semantik der Aussagenlogik**
  - Erweiterung der Interpretation I auf die Menge $\mathfrak{F}$ aller wohlgeformten Formeln
  - z.B. induktiv in Form von Wahrheitstabellen (siehe nächste Folie)

## Wahrheitstabelle

- **Seien $\phi$ und $\psi$ Formeln. Dann gilt:**

| $\phi$ | $\psi$ | $\neg\phi$ | $\phi \wedge \psi$ | $\phi \vee \psi$ | $\phi \Rightarrow \psi$ | $\phi \Leftrightarrow \psi$ |
|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| 0 | 0 | 1 | 0 | 0 | 1 | 1 |

- Eine Interpretation kann man als abstrakte Beschreibung der „möglichen Welt“ sehen
- Sie gibt an, welche Aussagen in dieser „Welt“ erfüllt sind und welche nicht erfüllt sind

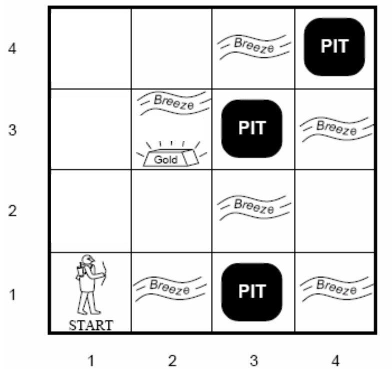
# Aussagenlogik

## Beispiel: Wissensbasis

- **Ziel:**
  - Konstruktion einer Wissensbasis für reduzierte Wumpus Welt
- **Einschränkung:**
  - Betrachte nur Fallgrube und Zugluft
  - Ignoriere Wumpus, Gestank, Gold und Geglitzer
- **Vokabular (Aussagensymbole):** Für alle i, j gelte
  - p[i,j] ist wahr, wenn im Raum (i,j) eine Fallgrube ist; sonst falsch
  - b[i,j] ist wahr, wenn im Raum (i,j) Zugluft ist; sonst falsch.

Wissensbasis für reduzierte Wumpus Welt (Startzustand)

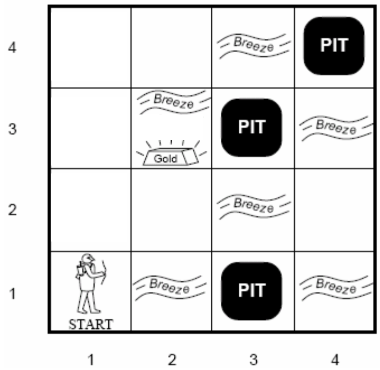
- Ein Raum windet gdw. im Nachbarfeld eine Grube ist
  - $\phi_1$ : b[1,1] $\Leftrightarrow$ (p[1,2] $\vee$ p[2,1])
  - $\phi_2$ : b[1,2] $\Leftrightarrow$ (p[1,1] $\vee$ p[1,3] $\vee$ p[2,2])
  - …
  - $\phi_{16}$: b[4,4] $\Leftrightarrow$ (p[4,3] $\vee$ p[3,4])
- Keine Fallgrube im Raum (1,1)
  - $\phi_{17}$: $\neg$p[1,1]
- Keine Zugluft im Raum (1,1)
  - $\phi_{18}$: $\neg$b[1,1]
- $\mathrm{WB}_0 = \phi_1 \wedge \ldots \wedge \phi_{18} = \{\phi_1, \ldots, \phi_{18}\}$



## Wissensbasis für reduzierte Wumpus Welt (Agent in (1,2))

- ⇨ **Wissensbasis aus Startzustand**
  - → $WB_0 = \phi_1 \land \ldots \land \phi_{18}$
- ⇨ **Zugluft im Raum** $(1,2)$
  - → $\phi_{19}$: b$[1,2]$
- ⇨ $WB_1 = \phi_1 \land \ldots \land \phi_{18} \land \phi_{19}$

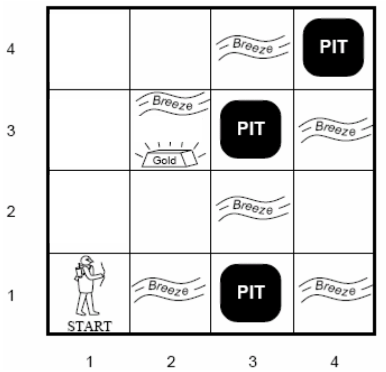
  $= \{\phi_1, \ldots, \phi_{19}\}$



# Gliederung

# Inferenz

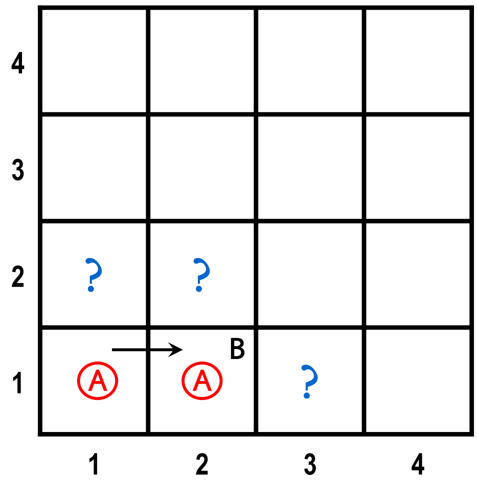
- ⇨ **Ausgangssituation:**
  - Agent repräsentiert Wissen in logischen Formeln. Um festzulegen, wie man neues Wissen folgern kann, muss formal ein Folgerungsbegriff definiert werden.
- ⇨ **Beispiel:**
  - Wie kann man aus $WB_0$ formal folgern, dass Feld (1,2) sicher ist?
- ⇨ **Eigenschaft des Folgerungsbegriffs:**
  - Wann immer WB wahr ist, dann sollte jede aus WB gefolgerte Formel ebenfalls wahr sein



## Inferenz

- **Modell**
  - **Def.:** Sei WB eine Menge von Formeln. Eine Interpretation **I** heißt Modell von WB, wenn **I** jeder Formel $\phi$ aus WB den Wahrheitswert $1$ zuweist.
- **Notationen:**
  - Ist I Modell von WB, dann sagen wir **I** erfüllt WB und schreiben **I** $\models$ WB
  - Gilt WB = $\{\phi\}$, dann schreiben wir **I** $\models \phi$ statt **I** $\models \{\phi\}$
- **Semantische Folgerung**
  - **Def.:** $\phi$ folgt semantisch (WB $\models \phi$) aus WB, wenn jedes Modell von WB auch Modell von $\phi$ ist.
  - Erfüllt die gewünschte Eigenschaft eines formalen Folgerungsbegriffs

- ⇨ **Ausgangssituation**
  - Wissensbasis $WB_1$
- ⇨ **Prüfe semantische Folgerungen**
  - $WB_1 \models \neg p[2,1]$ ?
    - Feld $(2,1)$ ist sicher?
  - $WB_1 \models \neg p[2,2]$ ?
    - Feld $(2,2)$ ist sicher?
- ⇨ **Inferenzmethode**
  - Model-Checking



- **Model-Checking zur Verifikation der Folgerbarkeit $\phi$ von aus $WB_1$**
  - Betrachte alle Interpretationen I von b[i,j], p[i,j]
  - Werte alle Formeln $\phi_k$ von $WB_1$ bzgl. Interpretationen I aus (Wahrheitstafel)
  - Werte Formel $\phi$ bzgl. Interpretationen I aus (Wahrheitstafel)
  - Identifiziere Menge $M(WB_1)$ und $M(\phi)$ der Modelle von $WB_1$ und $\phi$
  - $WB_1 \models \phi$, wenn $M(WB_1) \subseteq M(\phi)$
- **Ergebnis:**
  - 3 Interpretationen sind Modelle von $WB_1$
  - $WB_1 \models \neg p[2,1]$, d.h. $\neg p[2,1]$ ist wahr für jedes Modell von $WB_1$
    - Keine Fallgrube in Feld (2,1)
  - $\neg p[2,2]$ ist wahr für zwei und falsch für ein Modell von $WB_1$
    - $\neg p[2,2]$ folgt nicht aus $WB_1$
    - Feld (2,2) ist ein unsicheres Feld

## Inferenz

- **Model-Checking ist**
  - korrekt
    - d.h. es werden nur Formeln abgeleitet, die semantisch aus WB folgen
  - vollständig
    - d.h. es kann jede Formel abgeleitet werden, die semantisch aus WB folgt
  - exponentiell
    - $O(2^n)$ Interpretationen für n Symbole

- **Alternative:**
  - Resolutionsmethode

## Grundlagen

- **Erfüllbarkeit:**
  - $\phi$ heißt erfüllbar, wenn es ein Modell I von $\phi$ gibt
  - $\phi$ heißt unerfüllbar, wenn $\phi$ nicht erfüllbar ist
- **Gültigkeit:**
  - $\phi$ heißt gültig, wenn $\phi$ für alle Interpretationen I wahr ist
- **Äquivalenz:**
  - $\phi$ und $\psi$ heißen logisch äquivalent ($\phi \equiv \psi$), wenn für alle I gilt

$$\mathrm{I} \models \phi \quad \text{gdw.} \quad \mathrm{I} \models \psi$$

# Inferenz

- **Ziel:**
  - Prüfe ob WB $|= \phi$ gilt
- **Beispiel:**
  - WB $= (p \vee r) \wedge (q \vee \neg r)$ und $\phi = p \vee q$
  - $WB_1 |= \neg p[2,1]$
- **Resolution**
  - Methode zur Bestimmung der Unerfüllbarkeit einer Formel
  - Zeige WB $|= \phi$ durch Widerspruch: WB $\wedge \neg\phi$ ist unerfüllbar
  - Macht Sinn, weil (WB $|= \phi$) gdw. (WB $\wedge \neg\phi$) unerfüllbar (Beweis?)

## Informelle Beschreibung der Resolutionsmethode

1. Eingabe: Wissensbasis als Menge von Formeln
2. Wiederholte Anwendung einer syntaktische Umformungsregel
   a) Generiert in jedem Schritt aus zwei Formeln eine dritte Formel
   b) Fügt neu generierte Formel, sofern sie nicht in der Wissensbasis enthalten ist, als Eingabe für den nächsten Schritt hinzu
3. Terminiert, wenn ein Widerspruch oder keine neue Formel generiert werden kann, die nicht in der (erweiterten) Wissensbasis ist

⇨ **Resolutionsregel**

→ wird in Schritt 2a) als syntaktische Umformungsregel angewendet

⇨ **Widerspruch**

→ tritt auf, wenn eine unerfüllbare Formel generiert wird

## Resolutionsregel I:

$$(p \vee r) \wedge (q \vee \neg r) \models p \vee q$$

Formel 1

Formel 2

Aus Formeln 1 & 2 generierte Formel 3

## Beobachtung:

- Modell für $(p \vee r) \wedge (q \vee \neg r)$ ist auch Modell für $p \vee q$
- $(p \vee r) \wedge (q \vee \neg r)$ ist eine Konjunktion von Disjunktionen
- $r$ ist in beiden Disjunktionen komplementär
- **Problem:** In der Form nur selten anwendbar
- **Ziel:** Verallgemeinerung der Resolutionsregel I

## Grundlagen

- **Literale:**
  - Symbole oder negierte Symbole
  - Beispiel: $p, \neg p, q, \neg q, r, \neg r, \ldots$
- **Klausel:**
  - Disjunktion von Literalen
  - Beispiel: $(p \vee q \vee \neg r), (\neg p \vee \neg q), (p \vee q)$
- **Konjunktive Normalform (KNF):**
  - Konjunktion von Klauseln
  - Beispiel: $(p \vee r) \wedge (q \vee \neg r)$

## Grundlagen (Forts.): Konjunktive Normalform

- ⇨ **Satz:** Jede Formel kann in eine äquivalente Formel in KNF umgeformt werden.
- ⇨ Es existiert ein konstruktiver Beweis
- ⇨ Hier: Nur Konstruktionsregeln + Beispiel

## Regeln zur Umformung in KNF

1. Eliminiere „⇔" mit Hilfe von
   - $\phi \Leftrightarrow \psi \equiv (\phi \Rightarrow \psi) \wedge (\psi \Rightarrow \phi)$
2. Eliminiere „⇒" mit Hilfe von
   - $\phi \Rightarrow \psi \equiv \neg\phi \vee \psi$
3. Ziehe „¬" nach innen mit Hilfe von
   - $\neg(\neg\phi) \equiv \phi$
   - $\neg(\phi \wedge \psi) \equiv \neg\phi \vee \neg\psi$ (De Morgan)
   - $\neg(\phi \vee \psi) \equiv \neg\phi \wedge \neg\psi$ (De Morgan)
4. Löse verschachtelte „∧" und „∨" Verknüpfungen auf mit Hilfe von
   - $\phi \wedge (\psi \vee \varphi) \equiv (\phi \wedge \psi) \vee (\phi \wedge \varphi)$ (Distributionsgesetz 1)
   - $\phi \vee (\psi \wedge \varphi) \equiv (\phi \vee \psi) \wedge (\phi \vee \varphi)$ (Distributionsgesetz 2)

- **Beispiel:** Transformation von $p \Leftrightarrow (q \vee r)$ in KNF
  1. Eliminiere „$\Leftrightarrow$“
     - → $(p \Rightarrow (q \vee r)) \wedge ((q \vee r) \Rightarrow p)$
  2. Eliminiere „$\Rightarrow$“
     - → $(\neg p \vee (q \vee r)) \wedge (\neg(q \vee r) \vee p)$
  3. Ziehe „$\neg$“ nach innen
     - → $(\neg p \vee q \vee r) \wedge ((\neg q \wedge \neg r) \vee p)$
  4. Löse verschachtelte „$\wedge$“ und „$\vee$“ Verknüpfungen auf
     - → $(\neg p \vee q \vee r) \wedge (\neg q \vee p) \wedge (\neg r \vee p)$
- **KNF:** $(\neg p \vee q \vee r) \wedge (\neg q \vee p) \wedge (\neg r \vee p)$

## ⇨ Resolutionsregel II:

- Seien $L_1, \ldots, L_n$ und $L^*_1, \ldots, L^*_m$ Literale mit $L = L_i = \neg L^*_j$ für ein i mit $1 \leq i \leq n$ und ein j mit $1 \leq j \leq m$. Dann gilt:

$$(L_1 \vee \ldots \vee L_{i-1} \vee L \vee L_{i+1} \vee \ldots \vee L_n) \wedge (L^*_1 \vee \ldots \vee L^*_{j-1} \vee \neg L \vee L^*_{j+1} \vee \ldots \vee L^*_n)$$

$$\models$$

$$(L_1 \vee \ldots \vee L_{i-1} \vee L_{i+1} \vee \ldots \vee L_n) \vee (L^*_1 \vee \ldots \vee L^*_{i-1} \vee L^*_{i+1} \vee \ldots \vee L^*_n)$$

Resolvente

**Idee des Resolutionsbeweises:** Was nützt uns die Resolutionsregel?

- **Beobachtung:**
  - **Es sei $\phi = K_1 \wedge \ldots \wedge K_n$ eine Formel in KNF mit Klauseln $K_i$ und L ein Literal. Gibt es Klauseln $K_r = L$ und $K_s = \neg L$, dann ist $\phi$ unerfüllbar.**
    - Es kann nicht gleichzeitig **L** und **¬L** gelten
    - Damit ist $K_1 \wedge \ldots \wedge L \wedge \ldots \wedge \neg L \wedge \ldots \wedge K_n$ erst recht unerfüllbar
  - **Resolution mit L und ¬L liefert als Resolvente die leere Klausel ●**
- **Folgerung:**
  - **Versuche aus** WB $\wedge \neg\phi$ **die leere Klausel ● abzuleiten**

## ⇨ Resolutionsmethode für WB $\models \phi$

1. Transformiere WB $\wedge \neg\phi$ in KNF; Sei M die KNF.
2. Wiederhole folgende Schritte
   a) Wähle zwei Klauseln K und K* aus M, für die die Resolutionsregel anwendbar ist
   b) Erzeuge eine Resolvente R von K und K* durch Anwendung der Resolutionsregel
   c) Füge Klausel R in M hinzu, falls nicht vorhanden

   bis einer der folgenden Fälle eintritt

   → Fall 1: Die leere Klausel ● wird hergeleitet

   → Fall 2: Es können keine neuen Resolventen mehr gebildet werden

- **Ergebnis der Resolutionsmethode**
  - **Fall 1:** Die leere Klausel ● taucht als Resolvente auf. Folglich gilt:

    $$WB \wedge \neg\phi \text{ unerfüllbar und damit } WB \models \phi$$

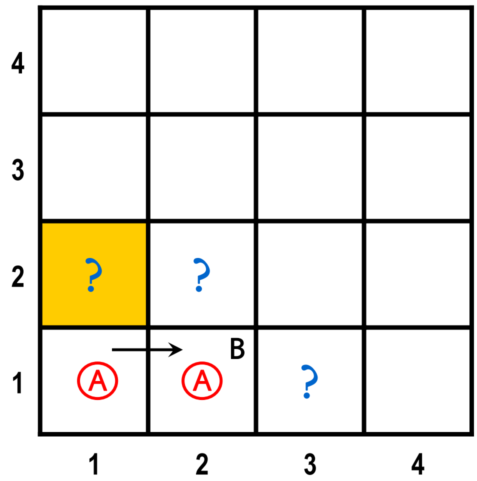
  - **Fall 2:** Die leere Klausel ● taucht nicht als Resolvente auf. Dann gilt:

    $$WB \wedge \neg\phi \text{ erfüllbar und damit } WB \not\models \phi$$

- **Eigenschaften der Resolutionsmethode**
  - **korrekt**
    - d.h. es werden nur Formeln abgeleitet, die semantisch aus WB folgen
  - **vollständig**
    - d.h. es kann jede Formel abgeleitet werden, die semantisch aus WB folgt
  - Findet Beweise,
    - d.h. Sequenz von Inferenzregeln, die aus WB eine Formel ableiten

## Inferenz

- **Ausgangssituation**
  - Wissensbasis $WB_1$
- **Prüfe semantische Folgerungen**
  - $WB_1 \models \neg p[2,1]$ ?
    - Feld (2,1) ist sicher?
- **Inferenzmethode**
  - Resolution
- **Relevante Wissensbasis**
  - $\phi_1$ : $b[1,1] \Leftrightarrow (p[1,2] \vee p[2,1])$
  - $\phi_{18}$: $\neg b[1,1]$
  - $WB_1 = \phi_1 \wedge \phi_{18}$



## Resolution

Beispiel

- **Ziel:**
  - Prüfe: $WB_1 \models \neg p[2,1]$ mit Resolution
- **Lösung**
  1. Transformiere $WB_1 \wedge p[2,1]$ in KNF
     - Ergebnis: (Klauselmenge: Konjunktionen als Kommata)

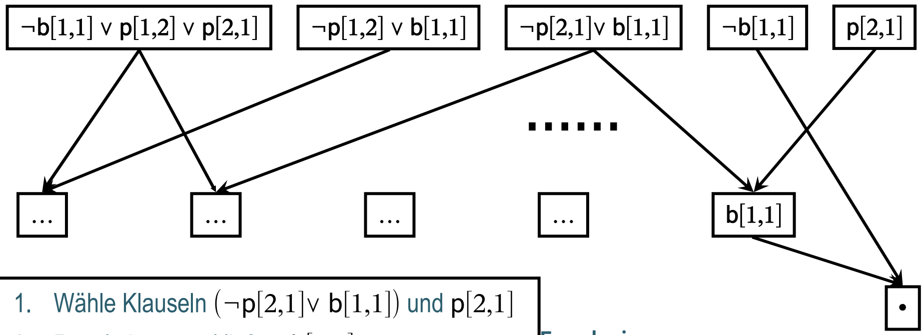
$$M = \{ (\neg b[1,1] \vee p[1,2] \vee p[2,1]), (\neg p[1,2] \vee b[1,1]), (\neg p[2,1] \vee b[1,1]), \neg b[1,1], p[2,1] \}$$

     - Siehe auch Folie #46 ´

## Resolution



### 1. Resolution:




1. Wähle Klauseln $(\neg p[2,1] \vee b[1,1])$ und $p[2,1]$
2. Resolutionsregel liefert: $b[1,1]$
3. Setze $M = M \cup \{b[1,1]\}$
4. Wähle Klauseln $b[1,1]$ und $\neg b[1,1]$
5. Resolutionsregel liefert leere Klausel

Ergebnis:

- → $WB_1 \wedge p[2,1]$ ist unerfüllbar
- → Es folgt $\neg p[2,1]$ aus $WB_1$
- → Feld $(2,1)$ ist sicher

## Gliederung

## Zusammenfassung & Ausblick

- **Zusammenfassung**
  - Intelligente Agenten benötigen eine Wissensbasis und einen Inferenzmechanismus um rationale Entscheidungen treffen zu können
  - Repräsentation der Wissensbasis durch Formeln der Aussagen-logik
  - Resolutionsmethode basierend auf KNF als Inferenzmechanis-mus
- **Ausblick**
  - Aussagenlogik sehr ausdruckschwach
  - Nächste Vorlesung: Prädikatenlogik der ersten Stufe

$$
\begin{aligned}
(\alpha \land \beta) &\equiv (\beta \land \alpha) && \text{commutativity of } \land \\
(\alpha \lor \beta) &\equiv (\beta \lor \alpha) && \text{commutativity of } \lor \\
((\alpha \land \beta) \land \gamma) &\equiv (\alpha \land (\beta \land \gamma)) && \text{associativity of } \land \\
((\alpha \lor \beta) \lor \gamma) &\equiv (\alpha \lor (\beta \lor \gamma)) && \text{associativity of } \lor \\
\neg(\neg\alpha) &\equiv \alpha && \text{double-negation elimination} \\
(\alpha \Rightarrow \beta) &\equiv (\neg\beta \Rightarrow \neg\alpha) && \text{contraposition} \\
(\alpha \Rightarrow \beta) &\equiv (\neg\alpha \lor \beta) && \text{implication elimination} \\
(\alpha \Leftrightarrow \beta) &\equiv ((\alpha \Rightarrow \beta) \land (\beta \Rightarrow \alpha)) && \text{biconditional elimination} \\
\neg(\alpha \land \beta) &\equiv (\neg\alpha \lor \neg\beta) && \text{de Morgan} \\
\neg(\alpha \lor \beta) &\equiv (\neg\alpha \land \neg\beta) && \text{de Morgan} \\
(\alpha \land (\beta \lor \gamma)) &\equiv ((\alpha \land \beta) \lor (\alpha \land \gamma)) && \text{distributivity of } \land \text{ over } \lor \\
(\alpha \lor (\beta \land \gamma)) &\equiv ((\alpha \lor \beta) \land (\alpha \lor \gamma)) && \text{distributivity of } \lor \text{ over } \land
\end{aligned}
$$