

# Einführung in das induktive Lernen

Prof. Dr.-Ing. R. Dillmann

Prof. Dr.-Ing. J. Marius Zöllner



Forschungszentrum Karlsruhe  
in der Helmholtz-Gemeinschaft



Universität Karlsruhe (TH)  
Forschungsuniversität • gegründet 1825

- Organisatorisches
- Induktion & Deduktion (4 - 9)
- Konzeptlernen als Suche im Hypothesenraum (10 - 18)
- Specific-to-general-Suche (19 - 29)
- Versionsraum (Version Space) / Candidate-Elimination-Algorithmus (30 - 44)
- Notwendigkeit von Vorzugskriterien (Bias) (45 - 50)

- Inhaltliche Fragen an jeweiligen Vortragenden
- Prüfungsvereinbarungen über Sekretariat!
  - Goettl@fzi.de
  - sekretariat.dillmann@anthropomatik.kit.edu
- Aktuellste Foliensätze nach der Vorlesung auf der ILIAS-Plattform

# Was ist Induktion?



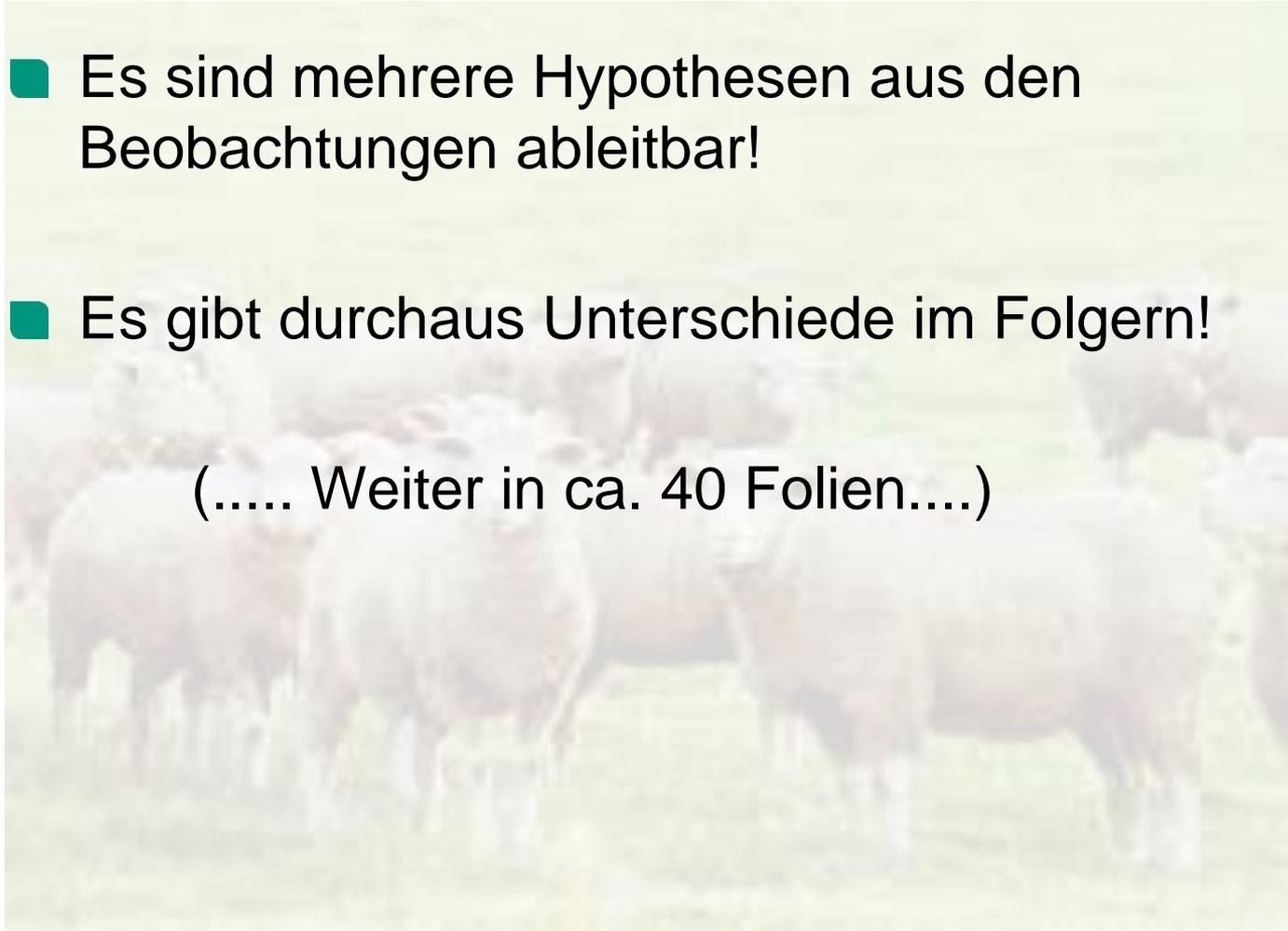
# Was ist Induktion?

- Ein Soziologe, ein Ingenieur, ein Physiker und ein Mathematiker fahren mit dem Zug in ein fernes Land. Kurz nachdem sie die Grenze passiert haben, sehen sie ein schwarzes Schaf.
- Meint der Soziologe: "Oh, wie interessant, ein schwarzes Schaf!"
- Darauf der Ingenieur: "Wir können jetzt annehmen, dass alle Schafe in diesem Land schwarz sind."
- Der Physiker darauf: "Nein. Wir können lediglich behaupten, dass ein Schaf in diesem Land schwarz ist."
- Der Mathematiker daraufhin: "Auch das ist falsch. Wir können lediglich sagen, dass es in diesem Land ein Schaf gibt, das von mindestens einer Seite schwarz ist."

# Was ist Induktion?

- Es sind mehrere Hypothesen aus den Beobachtungen ableitbar!
- Es gibt durchaus Unterschiede im Folgern!

(..... Weiter in ca. 40 Folien.....)



# Was ist Induktion?

- Prozess des **plausiblen** Schließens vom Speziellen zum Allgemeinen.
  - Basis: große Anzahl zutreffender Fälle.

$$P(x_1) \rightarrow Q(x_1)$$

$$P(x_2) \rightarrow Q(x_2)$$

$$\vdots \rightarrow \vdots$$

$$P(x_n) \rightarrow Q(x_n)$$

---

$$P(X) \rightarrow Q(X)$$

- **Aristoteles (384-324 v. Chr. , Logik):**
  - Induktion ist der Fortschritt vom Einzelnen zum Allgemeinen
- **Francis Bacon (1561-1626, Empirismus):**
  - Induktion ist das methodisch-experimentelle Sammeln und Bewerten von Beobachtungen, um dann Schritt für Schritt diese als Naturgesetze zu verallgemeinern.
- **John Stuart Mill (1806-1873, Philosoph u. Ökonom):**
  - Induktion ist das Analysieren von häufig wiederkehrenden Erfahrungen und das daraus folgende Schließen auf allgemeingültige Gesetzmäßigkeiten.
- **Karl R. Popper (1902-1994)**
  - Kritisiert in seiner Wissenschaftstheorie den reinen Ansatz der Induktion. Induktion mache nur Sinn in Zusammenhang mit einer deduktiven Überprüfung.

## ■ Induktion

- Wahrheitserweiternd
- Macht Lebewesen überlebensfähig
- Plausibilität

## ■ Deduktion

- Wahrheitserhaltend
- Logischer Schluss
- Korrektheit

- Formale Betrachtung eines induktiven Lernverfahrens:
  
- Gegeben:
  - Instanzraum  $X$
  - Trainingsmenge:  $D = x_1, \dots, x_n \in X$
  - Zielkonzept:  $c(..)$
  - Hypothesenraum  $H$
  
- Gesucht:
  - Hypothese  $h \in H$  mit  $h(x_i) = c(x_i)$ ,  $x_i \in X$

# Die induktive Lernhypothese

## ■ Induktive Lernhypothese:

Jede Hypothese, die die Zielfunktion über einer genügend großen Menge von Trainingsbeispielen gut genug approximiert, wird die Zielfunktion auch über unbekanntem Beispielen gut approximieren.

# Definition: Konzept

## ■ Konzept:

- Beschreibt Untermenge von Objekten oder Ereignissen definiert auf größerer Menge
- Bool'sche Funktion definiert über größerer Menge

## ■ Beispiel:

$\text{vogel} : \text{Tiere} \rightarrow \{ \text{true}, \text{false} \}$

$\text{vogel}(\text{Storch}) = \text{true}$

$\text{vogel}(\text{Hase}) = \text{false}$

## ■ Gegeben:

- Beispiele, die als Mitglieder oder Nichtmitglieder eines Konzepts gekennzeichnet sind

## ■ Gesucht:

- Automatischer Schluss auf die Definition des zugrundeliegenden Konzepts

## ■ Definition Konzeptlernen:

- Schließen auf eine Boolean-wertige Funktion aus Trainingsbeispielen ihres Inputs und Outputs. ([1])

## ■ Welches Konzept ist gemeint?

$c(\text{Wildkatze}) \rightarrow \text{true}$

$c(\text{Löwe}) \rightarrow \text{true}$

$c(\text{Leopard}) \rightarrow \text{true}$

$c(\text{Jaguar}) \rightarrow \text{false}$

$c(\text{Tiger}) \rightarrow \text{false}$

$c(\text{Zebra}) \rightarrow \text{true}$

$c(\text{Giraffe}) \rightarrow \text{true}$

$c(\text{Elefant}) \rightarrow \text{true}$

Gesucht: Tiere, die in Afrika heimisch sind

# Konzeptlernen: Ein Beispiel I

- (aus: [1])
- Zielkonzept:
  - „Tage an denen Aldo gerne Wassersport betreibt“
  - c: EnjoySport:  $X \rightarrow \{\text{true}, \text{false}\}$
- Gegeben 6 Attribute:
  - Himmel (sonnig, wolkig, regnerisch)
  - Lufttemperatur (warm, kalt)
  - Luftfeuchtigkeit (normal, hoch)
  - Wind (stark, schwach)
  - Wassertemperatur (warm, kalt)
  - Wettervorhersage (gleichbleibend, wechselnd)



- Menge der Instanzen X:
  - Mögliche Tage, beschrieben durch die 6 Attribute
    - <Himmel, Lufttemperatur, Luftfeuchtigkeit, Wind, Wassertemperatur, Wettervorhersage>
  - Bsp.: <sonnig, warm, normal, stark, warm, gleichbl.>
- Menge der möglichen Hypothesen H:
  - Jede Hypothese beschrieben durch Konjunktion von Einschränkungen auf die Attribute:
    - ? : Jeder Wert ist akzeptabel
    - value: Ein bestimmter Attributwert ist notwendig
    - # : Kein Wert ist akzeptabel

# Konzeptlernen: Ein Beispiel III

## ■ Syntax:

- $\langle ?, \text{kalt}, \text{hoch}, ?, ?, ? \rangle$  ist wahr  $\Leftrightarrow$  Lufttemperatur ist kalt und Luftfeuchtigkeit ist hoch (während alle übrigen Attribute beliebige Werte annehmen können)

## ■ Trainingsbeispiele:

| Himmel    | Lufttemp. | Luftfeuch. | Wind  | Wassert. | Vorhers.  | $c(x_i)$ |
|-----------|-----------|------------|-------|----------|-----------|----------|
| sonnig    | warm      | normal     | stark | warm     | gleichbl. | true     |
| sonnig    | warm      | hoch       | stark | warm     | gleichbl. | true     |
| regnersch | kalt      | hoch       | stark | warm     | wechselh. | false    |
| sonnig    | warm      | hoch       | stark | kalt     | wechselh. | true     |

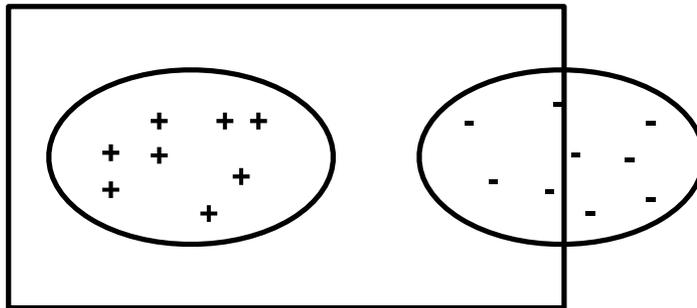
## ■ Zu bestimmen:

- Eine Hypothese  $h$  aus  $H$ , so dass  $h(x) = c(x) \forall x \in X$

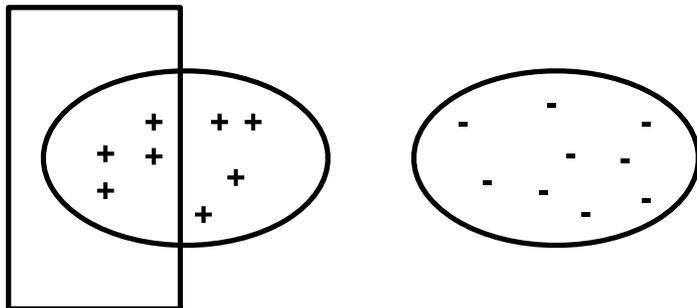
# Konsistenz und Vollständigkeit im Hypothesenraum - Definition

- **Konsistenz:** Keine negativen Beispiele werden positiv klassifiziert.
- **Vollständigkeit:** Alle positiven Beispiele werden als positiv klassifiziert.
- **Achtung:** Definition in [1] etwas anders, *konsistent* dort entspricht bei uns *konsistent + vollständig*

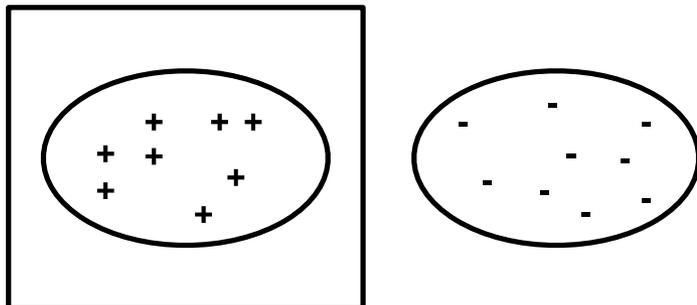
# Konsistenz und Vollständigkeit im Hypothesenraum - Beispiele



vollständig, aber  
nicht konsistent



konsistent, aber  
nicht vollständig

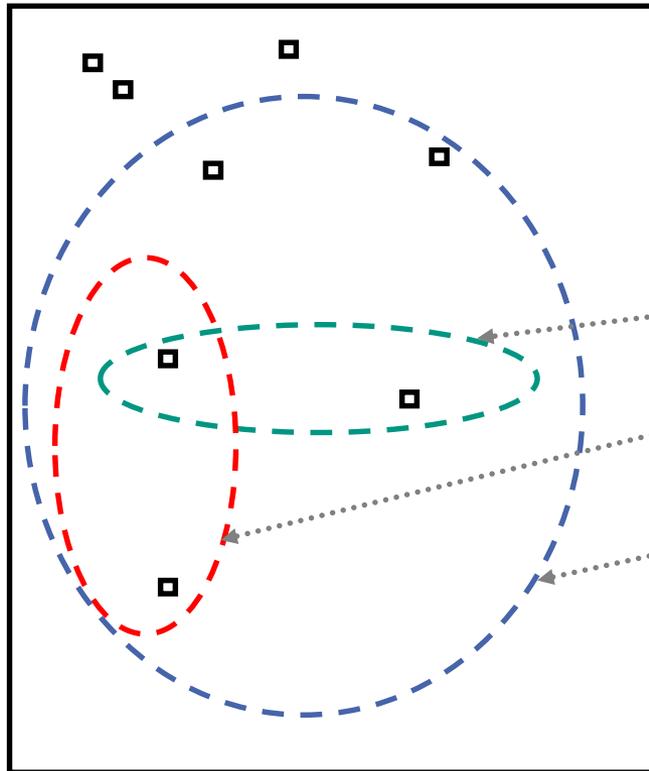


konsistent und vollständig  
(-> Vorzugskriterium)

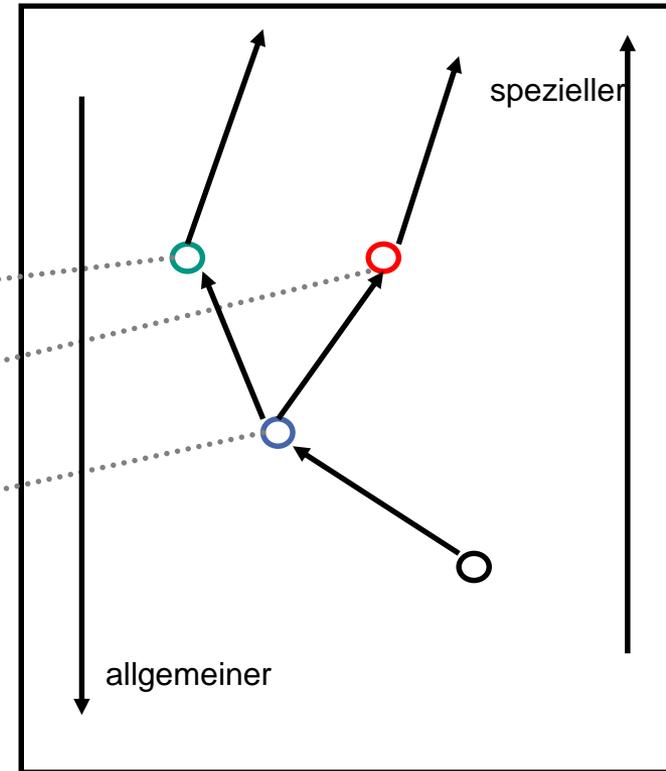
# Lernen als Suche im Hypothesenraum (1)

- Repräsentation der Hypothesen legt implizit Hypothesenraum fest (Domänenwissen → Bias)
- Lernen als Suche im Raum der möglichen Hypothesen:
  - 96 mögl. Instanzen, 973 mögl. Hypothesen
- Jeder Beschreibungsraum für Konzepte ist nach Generalität (halb-)geordnet:
  - $h_1 = \langle \text{sonnig, ?, ?, stark, ?, ?} \rangle$
  - $h_2 = \langle \text{sonnig, ?, ?, ?, ?, ?} \rangle$
  - $h_k$  spezieller  $h_i \Leftrightarrow \forall x \in X: [h_k(x) = 1 \Rightarrow h_i(x) = 1]$
  - wobei  $h(x) = 1$  bedeutet: *x erfüllt die Hypothese h*

# Hypothesenraum



Instanzen



Hypothese

<sonnig, ?, ?, ?, ?, ?>

<sonnig, ?, ?, stark, ?, ?>

<sonnig, ?, ?, ?, kühl, ?>

# Lernen als Suche im Hypothesenraum (2)

## ■ Suche vom Allgemeinen zum Speziellen:

- Ausgangspunkt ist allgemeinste Hypothese  $\langle ?, \dots, ? \rangle$
- Negative Beispiele: Spezialisierung
- Positive Beispiele: werden nicht betrachtet

## ■ Suche vom Speziellen zum Allgemeinen:

- Ausgangspunkt ist speziellste Hypothese  $\langle \#, \dots, \# \rangle$
- Positive Beispiele: (minimale) Verallgemeinerung
- Negative Beispiele: werden nicht betrachtet

## ■ Paralleles Anwenden beider Methoden $\Rightarrow$ Version Space

# Specific-to-General-Suche

- Initialisiere  $h$  mit der spezifischsten Hypothese in  $H$
- Für jedes positive Trainingsbeispiel  $x$ 
  - Für jede Attributeinschränkung  $a_i$  in  $h = \langle a_0, \dots, a_n \rangle$ 
    - Wenn  $a_i$  von  $x$  erfüllt wird
      - Dann tue nichts
    - Sonst:
      - Ersetze  $a_i$  durch die nächstallgemeinere Einschränkung, die durch  $x$  erfüllt wird
- Gib die Hypothese aus

# Specific-to-general Beispiel I

| Himmel     | Lufttemp. | Luftfeuch. | Wind  | Wassert. | Vorhers.  | $c(x_i)$ |
|------------|-----------|------------|-------|----------|-----------|----------|
| sonnig     | warm      | normal     | stark | warm     | gleichbl. | true     |
| sonnig     | warm      | hoch       | stark | warm     | gleichbl. | true     |
| regnerisch | kalt      | hoch       | stark | warm     | wechselh. | false    |
| sonnig     | warm      | hoch       | stark | kalt     | wechselh. | true     |

## ■ Initialisierung mit der **speziellsten Hypothese**

■  $h = \langle \#, \#, \#, \#, \#, \# \rangle$

## ■ 1. Beispiel: positiv

■  $x_1 = \langle \text{sonnig, warm, normal, stark, warm, gleichbl.} \rangle$

■  $c(x_1) = \text{true}$

■ Aber  $h(x_1) = \text{false} \Rightarrow$  Hypothese zu speziell, Verallgemeinerung

■  $h = \langle \text{sonnig, warm, normal, stark, warm, gleichbl.} \rangle$

# Specific-to-general Beispiel II

| Himmel     | Lufttemp. | Luftfeuch. | Wind  | Wassert. | Vorhers.  | $c(x_i)$ |
|------------|-----------|------------|-------|----------|-----------|----------|
| sonnig     | warm      | normal     | stark | warm     | gleichbl. | true     |
| sonnig     | warm      | hoch       | stark | warm     | gleichbl. | true     |
| regnerisch | kalt      | hoch       | stark | warm     | wechselh. | false    |
| sonnig     | warm      | hoch       | stark | kalt     | wechselh. | true     |

- $h = \langle \text{sonnig, warm, normal, stark, warm, gleichbl.} \rangle$
- 2. Beispiel: positiv
  - $x_2 = \langle \text{sonnig, warm, hoch, stark, warm, gleichbl.} \rangle$
  - $c(x_2) = \text{true}$
  - Aber nicht von  $h$  abgedeckt,  $h(x_2) = \text{false} \Rightarrow$  minimale Verallgemeinerung, so dass  $h(x_2) = \text{true}$
  - $h = \langle \text{sonnig, warm, ?, stark, warm, gleichbl.} \rangle$

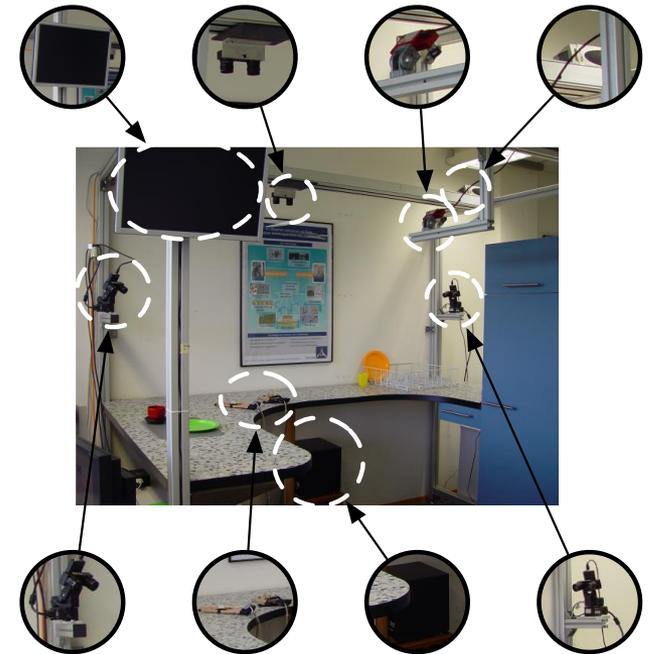
# Specific-to-general Beispiel III

| Himmel     | Lufttemp. | Luftfeuch. | Wind  | Wassert. | Vorhers.  | $c(x_i)$ |
|------------|-----------|------------|-------|----------|-----------|----------|
| sonnig     | warm      | normal     | stark | warm     | gleichbl. | true     |
| sonnig     | warm      | hoch       | stark | warm     | gleichbl. | true     |
| regnerisch | kalt      | hoch       | stark | warm     | wechselh. | false    |
| sonnig     | warm      | hoch       | stark | kalt     | wechselh. | true     |

- $h = \langle \text{sonnig, warm, ?, stark, warm, gleichbl.} \rangle$
- 3. Beispiel: negativ
  - Wird ignoriert
- 4. Beispiel
  - $x_4 = \langle \text{sonnig, warm, hoch, stark, kalt, wechselhaft} \rangle$
  - $c(x_4) = \text{true}$
  - Aber nicht von  $h$  abgedeckt,  $h(x_4) = \text{false} \Rightarrow$  minimale Verallgemeinerung, so dass  $h(x_4) = \text{true}$
  - $h = \langle \text{sonnig, warm, ?, stark, ?, ?} \rangle$

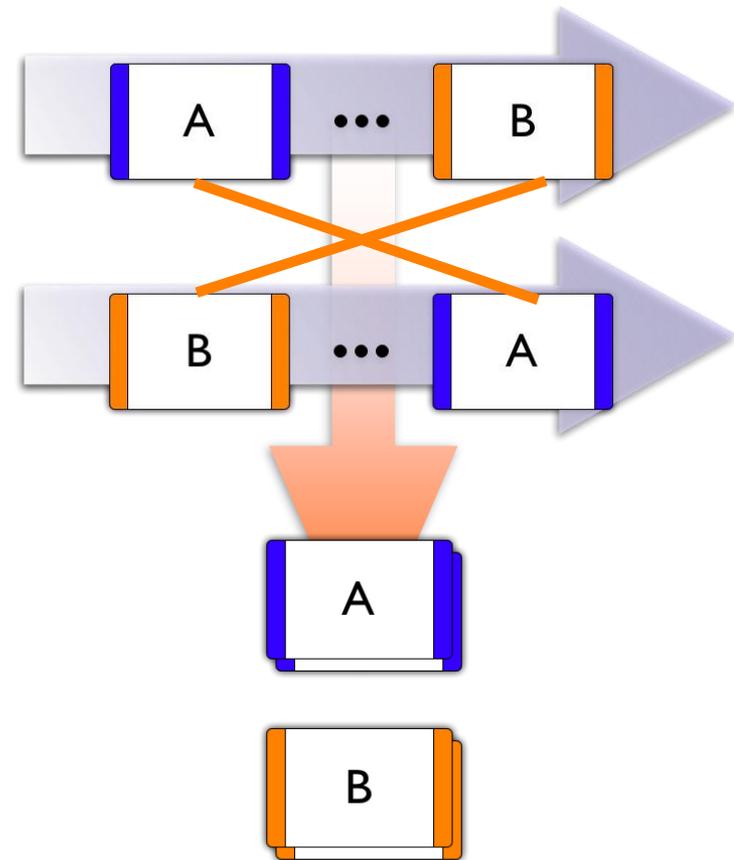
# Anwendung: Lernen von Präzedenzgraphen im PdV

- Programmieren von Servicerobotern durch Vormachen, im folgenden: Akquisition von Aufgabenwissen
- Intuitive Benutzerschnittstellen
  - Benutzer führt Aufgabe aus
  - Beobachten & Lernen durch PdV-System
  - Ausführung durch Roboter
- Umordnung von Teilaufgaben
  - Optimierung von Wegen, Energieeffizienz
  - Lernen der Vorrangbeziehungen zwischen Teilaufgaben



# Lernen sequentieller Unabhängigkeiten

- Mehrere Vorführungen
- Identifikation äquivalenter Teiloperatoren
- Operationen abhängig/unabhängig



# Definition: Präzedenzgraphen

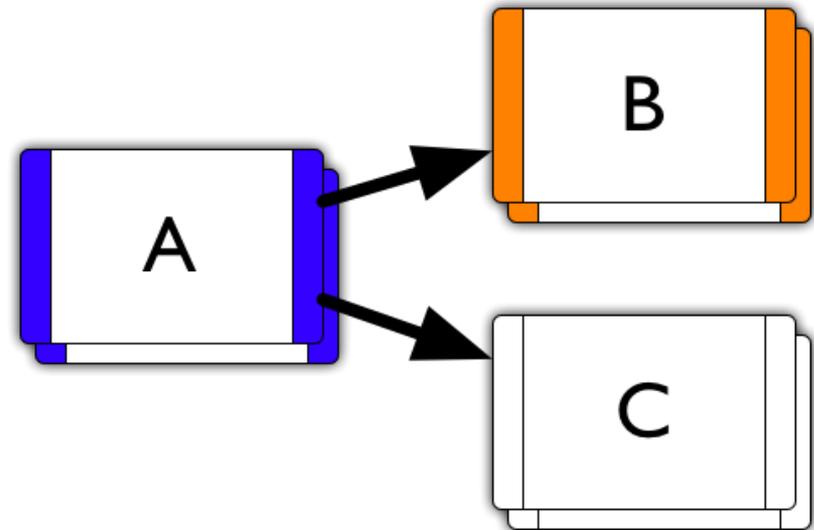
- Codierung der Hypothesen
- DAG (=Gerichteter azyklischer Graph)

$$P = (N, E)$$

- Kanten:  
Vorrangbeziehungen

$$(o_1, o_2) \in E$$

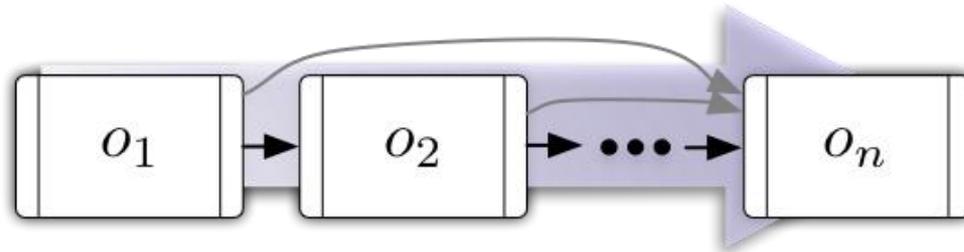
$$o_1, o_2 \in N$$



## ■ Hypothese zu einzelner Demonstration

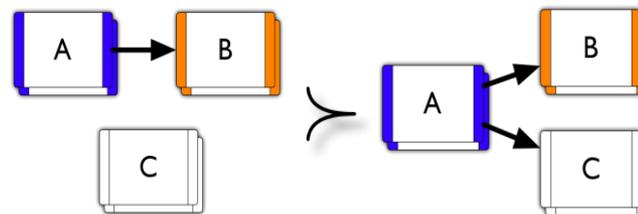
$$D = (o_1, o_2, \dots, o_n)$$

$$h = (N, E) \text{ mit } N = \{o_1, \dots, o_n\}, \forall j > i : (o_i, o_j) \in E$$

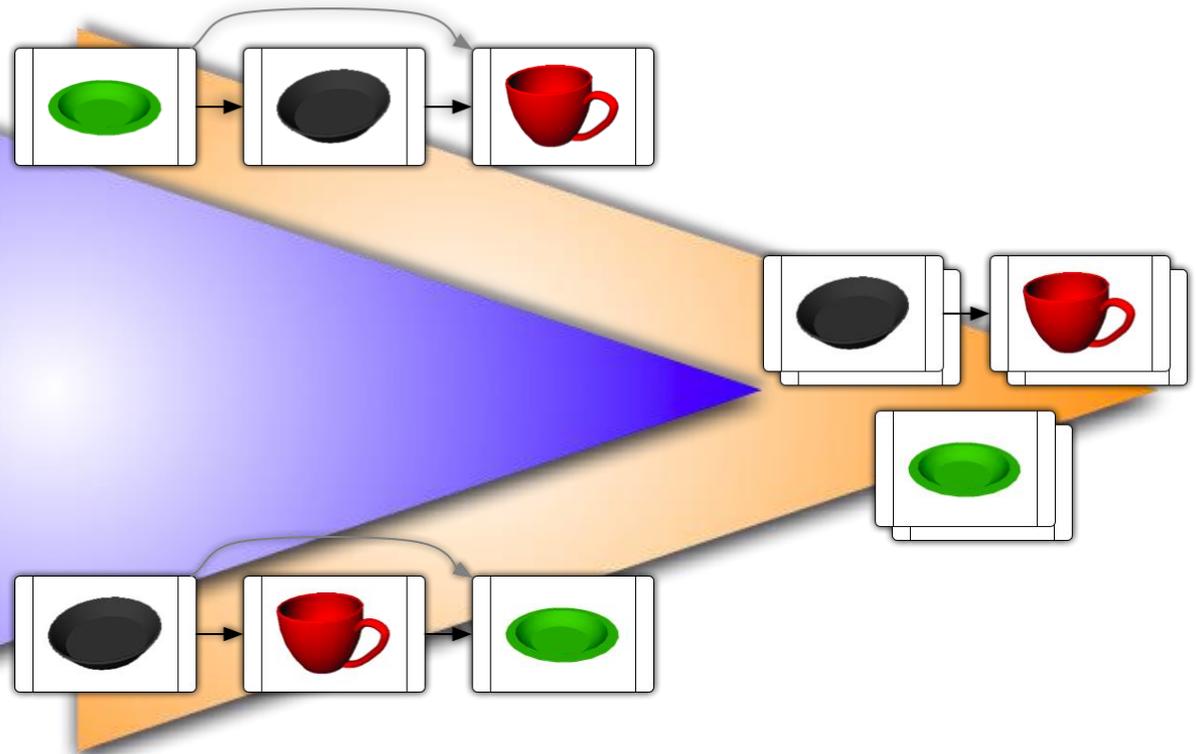
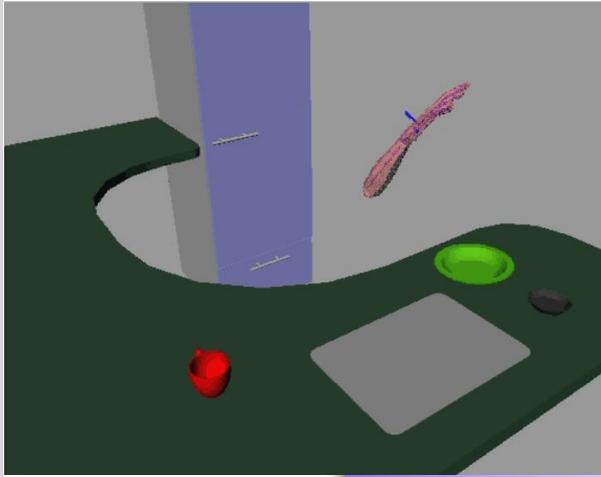


## ■ Genereller-als-Beziehung zwischen Präzedenzgraphen:

$$h_1 \succ h_2 \Leftrightarrow E_1 \subset E_2$$



# Lernen von Präzedenzgraphen II



# Specific to General: Beurteilung I

- Wichtiges Prinzip im Konzeptlernen
- Für Hypothesenräume, die durch Konjunktionen von Attributeinschränkungen beschrieben sind garantiert das Verfahren die spezifischste Hypothese, die mit den positiven Trainingsbeispielen vereinbar ist
- Endhypothese ist auch mit negativen Trainingsbeispielen konsistent, solange
  - die Trainingsbeispiele korrekt sind
  - die Zielhypothese in  $H$  enthalten ist

# Specific to General: Beurteilung II

- Offene Fragen:
  - Sind die Trainingsbeispiele konsistent?
  - Endkonzept = korrektes Zielkonzept?
  - Warum spezifischste Hypothese?

# Definition: Versionsraum

## ■ Definition: Versionsraum (Version Space)

Der Versionsraum  $VS_{H,D}$  bezüglich des Hypothesenraums  $H$  und der Menge von Trainingsbeispielen  $D$  ist die Untermenge der Hypothesen von  $H$ , die mit den Trainingsbeispielen in  $D$  konsistent ist.

# Version Space / Candidate-Elimination Algorithmus (Mitchell 1982)

- Lernen ist inkrementell:
  - Menge der konsistenten Hypothesen (Version space) ist ein Intervall in der partiellen Ordnung "spezifischer als" auf dem Hypothesenraum
- Gespeichert werden:
  - Menge der spezifischsten Hypothesen  $S$  und Menge der allgemeinsten Hypothesen  $G$ , die alle Beispiele abdecken  $\Rightarrow$  Hypothesen müssen nicht einzeln gespeichert werden
- Spezifischste und Allgemeinste Hypothesen:
  - $S = \{s \mid s \text{ ist eine Hypothese, die mit den betrachteten Beispielen konsistent ist, und es gibt keine Hypothese, die spezifischer als } s \text{ und auch konsistent mit allen Beispielen ist}\}$ .
  - Initialisierung:  $S = \{\#\}$
  - $G = \{g \mid g \text{ ist eine Hypothese, die mit den betrachteten Beispielen konsistent ist, und es gibt keine Hypothese, die allgemeiner als } g \text{ und auch konsistent mit allen Beispielen ist}\}$
  - Initialisierung:  $G = \{?\}$

## Ist $n$ ein negatives Beispiel:

- Lösche aus  $S$  die Hypothesen, die  $n$  abdecken.
- Spezialisierere die Hypothesen in  $G$  soweit, dass sie  $n$  nicht abdecken und dass sie allgemeiner als eine Hypothese in  $S$  bleiben.
- Lösche aus  $G$  alle Hypothesen, die spezifischer als eine andere Hypothese aus  $G$  sind.

## Ist $p$ ein positives Beispiel:

- Lösche aus  $G$  die mit  $p$  inkonsistenten Hypothesen.
- Verallgemeinere die Hypothesen in  $S$  soweit, dass sie  $p$  abdecken und dass sie spezifischer als eine Hypothese in  $G$  bleiben.
- Lösche aus  $S$  alle Hypothesen, die allgemeiner als eine andere Hypothese aus  $S$  sind



# Version-Space Beispiel I

## ■ Initialisierung:

■  $S_0 = \{ \langle \#, \#, \#, \#, \#, \# \rangle \}$

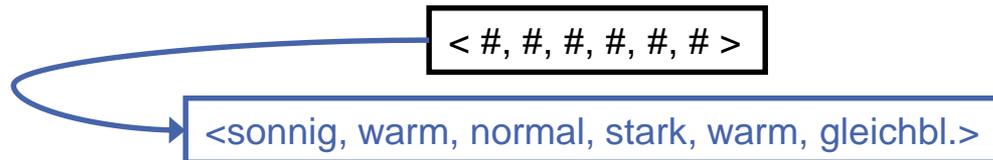
■  $G_0 = \{ \langle ?, ?, ?, ?, ?, ? \rangle \}$

$\langle \#, \#, \#, \#, \#, \# \rangle$

$\langle ?, ?, ?, ?, ?, ? \rangle$

# Version Space Beispiel II

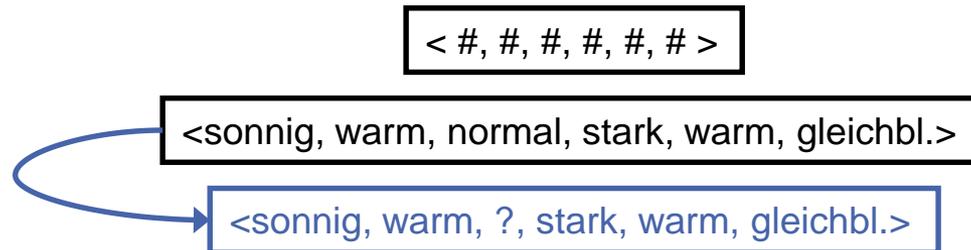
- $x_1 = \langle \text{sonnig, warm, normal, stark, warm, gleichbl.} \rangle$ ,  $c(x_1) = \text{true}$
- G konsistent mit  $x_1$
- S zu speziell  $\Rightarrow$  verallgemeinern bis konsistent



$\langle ?, ?, ?, ?, ?, ? \rangle$

# Version Space Beispiel III

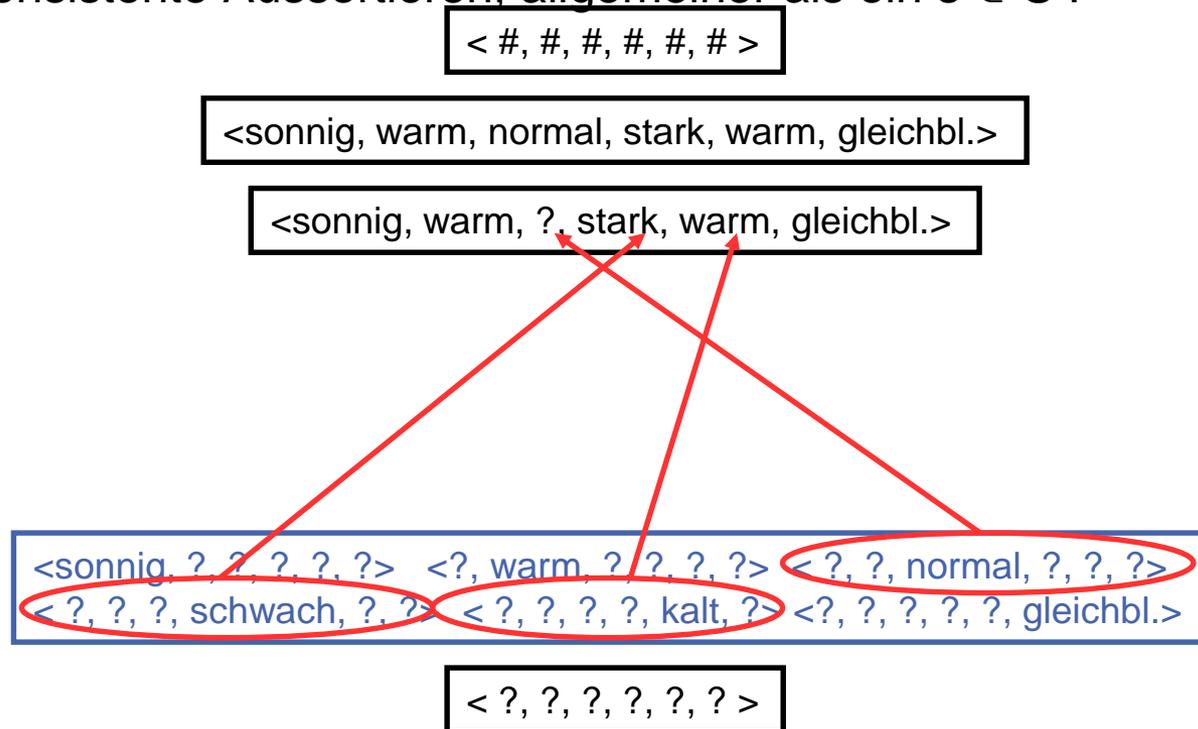
- $x_2 = \langle \text{sonnig, warm, hoch, stark, warm, gleichbl.} \rangle$ ,  $c(x_2) = \text{true}$
- G konsistent mit  $x_2$
- S zu speziell  $\Rightarrow$  verallgemeinern bis konsistent



$\langle ?, ?, ?, ?, ?, ? \rangle$

# Version Space Beispiel IV

- $x_3 = \langle \text{regnerisch, kalt, hoch, stark, warm, wechselh} \rangle$ ,  $c(x_3) = \text{false}$
- S konsistent mit  $x_3$
- G zu allgemein, suche minimale Spezialisierungen
  - Inkonsistente Aussortieren, allgemeiner als ein  $s \in S$  !



# Version Space Beispiel IV

- $x_3 = \langle \text{regnerisch, kalt, hoch, stark, warm, wechselh} \rangle$ ,  $c(x_3) = \text{false}$
- S konsistent mit  $x_3$
- G zu allgemein, suche minimale Spezialisierungen
  - Inkonsistente Aussortieren, allgemeiner als ein  $s \in S$  !

$\langle \#, \#, \#, \#, \#, \# \rangle$

$\langle \text{sonnig, warm, normal, stark, warm, gleichbl.} \rangle$

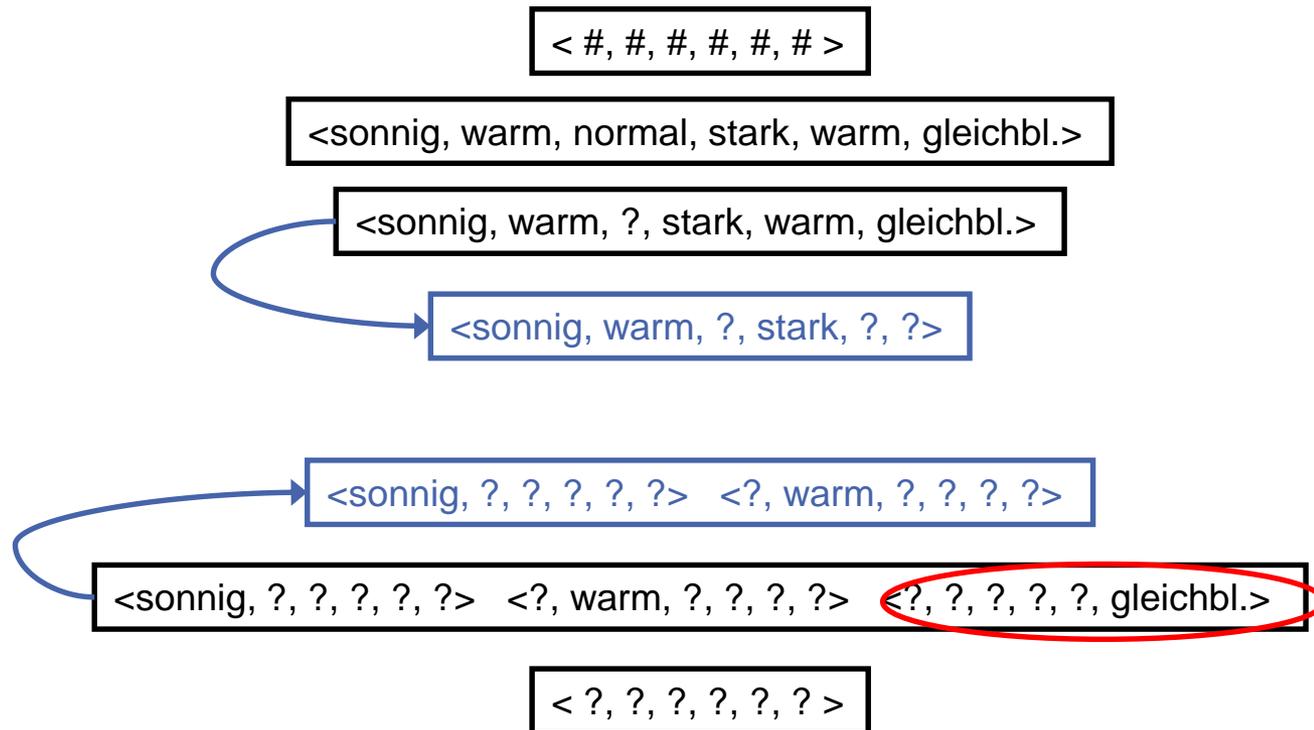
$\langle \text{sonnig, warm, ?, stark, warm, gleichbl.} \rangle$

$\langle \text{sonnig, ?, ?, ?, ?, ?} \rangle$   $\langle \text{?, warm, ?, ?, ?, ?} \rangle$   $\langle \text{?, ?, ?, ?, ?, gleichbl.} \rangle$

$\langle \text{?, ?, ?, ?, ?, ?} \rangle$

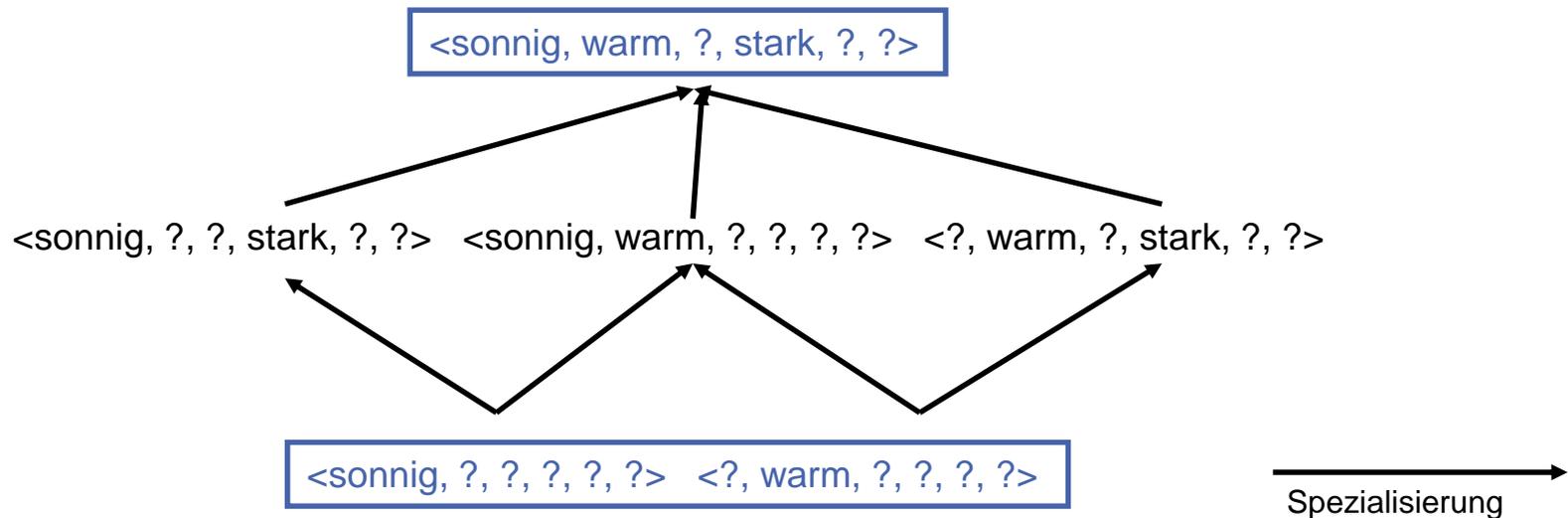
# Version Space Beispiel V

- $x_2 = \langle \text{sonnig, warm, hoch, stark, kühl, wechselh.} \rangle$ ,  $c(x_2) = \text{true}$
- Eine Hypothese in  $G$  inkonsistent mit  $x_2$
- $S$  nicht konsistent mit  $x_2 \Rightarrow$  verallgemeinern bis konsistent



# Version Space = alle konsistenten Hypothesen

| Himmel    | Lufttemp. | Luftfeuch. | Wind  | Wassert. | Vorhers.  | $c(x_i)$ |
|-----------|-----------|------------|-------|----------|-----------|----------|
| sonnig    | warm      | normal     | stark | warm     | gleichbl. | true     |
| sonnig    | warm      | hoch       | stark | warm     | gleichbl. | true     |
| regnersch | kalt      | hoch       | stark | warm     | wechselh. | false    |
| sonnig    | warm      | hoch       | stark | kalt     | wechselh. | true     |



# Einordnung

|                      |                     |   |                           |
|----------------------|---------------------|---|---------------------------|
| Typ der Inferenz     | <i>induktiv</i>     | ↔ | <i>deduktiv</i>           |
| Ebenen des Lernens   | <i>symbolisch</i>   | ↔ | <i>subsymbolisch</i>      |
| Lernvorgang          | <i>überwacht</i>    | ↔ | <i>unüberwacht</i>        |
| Beispielgebung       | <i>inkrementell</i> | ↔ | <i>nicht inkrementell</i> |
| Umfang der Beispiele | <i>umfangreich</i>  | ↔ | <i>gering</i>             |
| Hintergrundwissen    | <i>empirisch</i>    | ↔ | <i>axiomatisch</i>        |

- Version Space konvergiert zur korrekten Hypothese (S=G)
  - Voraussetzung:
    - Beispiele konsistent
    - Korrekte Hypothese in Hypothesenraum enthalten
  - Probleme:
    - fehlerbehaftete Trainingsdaten (Rauschen)!
    - Zielkonzept nicht von Hypothesenrepräsentation abgedeckt  
⇒ mögliche Erweiterung: disjunktive Begriffe
- Anfrage des Konzeptlernalers:
  - möglichst so, dass Hypothesen im Version Space halbiert werden
  - Dann: Schnelle Lernrate / geringe Anzahl von Beispielen

- Umgang mit teilweise gelernten Konzepten nötig

Wenn mehr als eine Hypothese im Versionsraum vorhanden ist dann:

- Alle klassifizieren positiv bzw. negativ  $\Rightarrow$  Entscheidung klar
- Ansonsten:
  - Mehrheitsentscheidung, evtl. mit Angabe von Konfidenz
  - Probabilistische Entscheidung / Wahrscheinlichkeit
  - Plausibilitätsbetrachtungen

# Version Space: Beurteilung III

- Konsistente Beispiele notwendig
- Attributgeneralisierungsregeln maßgebend für Lernerfolg
- + Kein Speichern alter Beispiele notwendig
- + Stellt fest, wann genügend Beispiele gegeben wurden (S=G)
- Unter Umständen Art noch benötigter Beispiele erkennbar

# Induktiver Bias



- Dem Soziologen wird es zu bunt, er zieht die Notbremse, der Zug kommt zum Stehen und die 4 steigen aus, um den Dingen auf den Grund zu gehen. Als sie das Tier erreicht haben, stellen sie fest, dass es tatsächlich auf der einen Seite weiß ist und auf der anderen Seite schwarz mit kleinen, aus der Ferne nicht erkennbaren weißen Flecken. Mittlerweile eilt der Bauer heran, der sich über den Aufmarsch auf seinem Feld wundert.
- Der Soziologe spricht ihn an: "Komische Schafe haben Sie hier!"
- Daraufhin der Bauer: „Das ist kein Schaf. Das ist eine Ziege.“

- Bisher ging es nur darum, wie sehr man aufgrund von Beispielen verallgemeinern kann.
- Der Raum der betrachteten Hypothesen spielt aber eine wesentliche Rolle !!!
- In diesem Fall z.B. war die ganze Diskussion bzgl. der Farbe „an der Sache vorbei“, da ALLE Hypothesen falsch waren, weil immer nur von Schafen die Rede war.

# Einschränkung des Hypothesenraums?

- Problem: Zielkonzept evtl. nicht im Hypothesenraum enthalten
- Lösung: Hypothesenraum, der alle möglichen Hypothesen enthält?
- Grundlegende Eigenschaft von induktiver Inferenz:
  - Ein induktives Lernsystem, das keine a priori-Annahmen über die Identität des Zielkonzepts macht, hat keine rationale Basis, um unbekannte Instanzen zu klassifizieren.
- Induktives Lernen erfordert Vorannahmen („inductive bias“)

# Vergleich induktiver Lernsysteme anhand ihres inductive bias

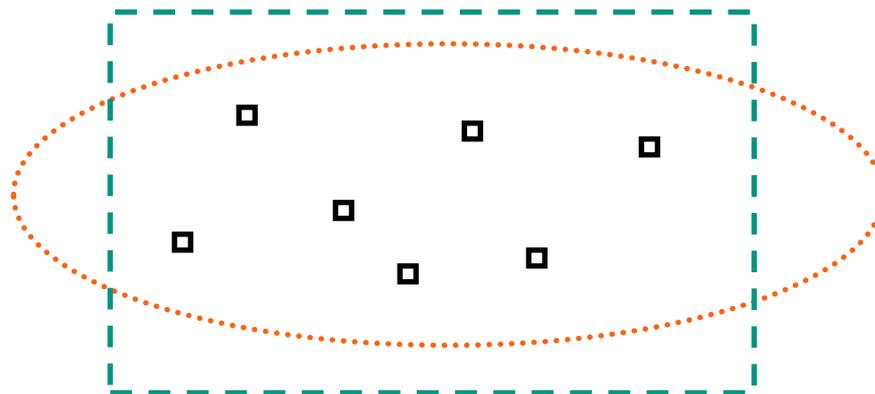
- **Auswendiglerner:**
  - keine Vorannahme
- **Version Space:**
  - Zielkonzept kann im Hypothesenraum repräsentiert werden
- **Specific to General-Lerner:**
  - wie Version Space, und zusätzlich:
    - alle Instanzen sind negative Instanzen, solange nicht das Gegenteil bekannt ist
- Je strenger die Vorannahmen, desto mehr unbekannte Beispiele können klassifiziert werden!

## ■ Bias (Vorzugskriterium)

- Vorschrift nach der Hypothesen gebildet werden.

## ■ Mögliche Vorzugskriterien:

- Verständlichkeit (für den menschlichen Benutzer)
- Klassifikationsgenauigkeit
- Messaufwand für die verwendeten Deskriptoren
- Berechnungs- und Speicheraufwand für die Hypothese



## ■ Hypothesenraumbias:

- $h$  gehöre zu einem beschränkten Raum von Hypothesen,
  - logische Konjunktionen
  - lineare Schwellwertfunktionen
  - Geraden, Polynome ...etc
  - 3-NN (Nearest Neighbour)
  - ...

## ■ Präferenzbias:

- Ordnung auf dem Raum der Hypothesen.
- Wähle  $h$  mit der höchsten Präferenz.
  - Bevorzuge Hypothesen mit weniger Disjunktionen
  - Bevorzuge kleinere Entscheidungsbäume
  - ....

## ■ Problem:

- Es existiert keine Funktion  $h$ , die konsistent mit allen Trainingsbeispielen ist, z.B. wegen verrauschter Trainingsdaten.

## ■ Unterschiedliche Ansätze möglich → unterschiedliche Lösungen:

## ■ Anpassen des Hypothesenraumbias:

- Problem, da zwar sehr gute Klassifikation i.a. durch eine komplexe Hypothese erreicht wird aber
- Overfitting !

## ■ Anpassen des Präferenzbias:

- Wähle das  $h \in H$ , das möglichst viele Beispiele richtig klassifiziert
- Misklassifikation muss in Kauf genommen werden

- Induktion & Deduktion
- Konzeptlernen als Suche im Hypothesenraum
- Specific to general-Suche
- Versionsraum (Version Space) / Candidate-Elimination-Algorithmus
- Notwendigkeit von Vorzugskriterien (Bias)

- [1] *Tom Mitchell: **Machine Learning***. McGraw-Hill, New York, 1997.
- [2] *Garry Briscoe, Terry Caeli: **A Compendium of Machine Learning, Vol.1: Symbolic Machine Learning***. Ablex Publishing, Norwood, NJ, 1996.
- [3] *Tom Mitchell: **Generalization as Search***. Artificial Intelligence, Vol. 18, 1982, S. 203-226