

Grundlagen der Künstlichen Intelligenz

12. Aussagenlogik: Grundlagen, Schliessen und Resolution

Malte Helmert

Universität Basel

22. April 2013

Grundlagen der Künstlichen Intelligenz

22. April 2013 — 12. Aussagenlogik: Grundlagen, Schliessen und Resolution

12.1 Motivation

12.2 Grundlagen

12.3 Logisches Schliessen

12.4 Resolution

12.1 Motivation

Aussagenlogik: Motivation

Diese Woche: Aussagenlogik

- ▶ Modellierung und Repräsentation von Problemen und Wissen
- ▶ Grundlage **allgemeiner** Problembeschreibungen und Lösungsstrategien
($\rightsquigarrow$ Handlungsplanung, nächste Woche)
- ▶ erlaubt automatisches **logisches Schliessen** (Thema heute)

Zusammenhang zu CSPs

- ▶ **letzte Woche:** Constraint-Satisfaction-Probleme
- ▶ Erfüllbarkeitsproblem der Aussagenlogik kann als **nicht-binäres CSP** über $\{\mathbf{F}, \mathbf{T}\}$ aufgefasst werden
- ▶ Formel kodiert die Constraints
- ▶ Gesucht: erfüllende Zuweisung von Werten zu Variablen
- ▶ SAT-Algorithmen für diese Probleme: $\rightsquigarrow$ DPLL (**Freitag**)

Aussagenlogik: Beschreibung von Zustandsräumen

Aussagevariablen für Missionare und Kannibalen:

```
two-missionaries-on-left-shore  
one-cannibal-on-left-shore  
boat-on-left-shore  
...
```

- ▶ Problembeschreibung verständlich für allgemeine Problemlöser
- ▶ Zustände repräsentiert als Wahrheitswerte der atomaren **Propositionen** (Aussagevariablen)

12.2 Grundlagen

Aussagenlogik: Intuition

Propositionen: atomare (nicht weiter teilbare)
Aussagen über die Welt.

Propositionen bilden mit **logischen Verknüpfungen**
wie „und“, „oder“ und „nicht“ die aussagenlogischen Formeln.

Syntax

Σ Alphabet von Propositionen
(z. B. $\{P, Q, R, \dots\}$ oder $\{X_1, X_2, X_3, \dots\}$).

Definition (aussagenlogische Formeln)

- ▶ $\top$ und $\perp$ sind Formeln.
- ▶ Jede Proposition aus Σ ist eine (atomare) Formel.
- ▶ Wenn φ eine Formel ist, dann auch $\neg\varphi$ (Negation).

Wenn φ und ψ Formeln sind, dann auch

- ▶ $\varphi \wedge \psi$ (Konjunktion)
- ▶ $\varphi \vee \psi$ (Disjunktion)
- ▶ $\varphi \rightarrow \psi$ (Implikation)

Bindungsstärke: $(\neg) > (\wedge) > (\vee) > (\rightarrow)$
(bei Bedarf Klammern setzen)

Semantik

Eine Formel ist **wahr** oder **falsch**; je nach dem, wie ihre Propositionen **interpretiert** werden.

Semantik: Intuition

- ▶ Propositionen sind entweder wahr oder falsch. Dies wird durch eine **Interpretation** bestimmt.
- ▶ Wahrheitswert einer Formel folgt aus Wahrheitswerten ihrer Propositionen.

Beispiel

$$\varphi = (P \vee Q) \wedge R$$

Wahrheitswert für beispielhafte Interpretationen:

- ▶ Wenn P und Q falsch sind und R wahr ist, dann ist φ falsch.
- ▶ Wenn P und R wahr sind, dann ist auch φ wahr (unabhängig vom Wahrheitswert von Q).

Semantik: formal

- ▶ definiert über **Interpretation** $I : \Sigma \rightarrow \{\mathbf{T}, \mathbf{F}\}$
- ▶ Interpretation I ist **Belegung** der Propositionen in Σ .
- ▶ Wann ist eine Formel φ unter einer Interpretation I wahr?
Symbolisch: Wann gilt $I \models \varphi$?

Definition ($I \models \varphi$)

- ▶ $I \models \top$ und $I \not\models \perp$
- ▶ $I \models P$ gdw. $I(P) = \mathbf{T}$ für $P \in \Sigma$
- ▶ $I \models \neg\varphi$ gdw. $I \not\models \varphi$
- ▶ $I \models \varphi \wedge \psi$ gdw. $I \models \varphi$ und $I \models \psi$
- ▶ $I \models \varphi \vee \psi$ gdw. $I \models \varphi$ oder $I \models \psi$
- ▶ $I \models \varphi \rightarrow \psi$ gdw. $I \not\models \varphi$ oder $I \models \psi$
- ▶ $I \models \Phi$ für eine Formelmeng Φ gdw. $I \models \varphi$ für alle $\varphi \in \Phi$

Beispiele

Beispiel (Interpretation I)

- ▶ $I(P) = \mathbf{T}$
- ▶ $I(Q) = \mathbf{T}$
- ▶ $I(R) = \mathbf{F}$
- ▶ $I(S) = \mathbf{F}$

Welche Formeln sind unter I wahr?

- ▶ $\varphi_1 = \neg(P \wedge Q) \wedge (R \wedge \neg S)$. Gilt $I \models \varphi_1$?
- ▶ $\varphi_2 = (P \wedge Q) \wedge \neg(R \wedge \neg S)$. Gilt $I \models \varphi_2$?
- ▶ $\varphi_3 = (R \rightarrow P)$. Gilt $I \models \varphi_3$?

Terminologie

Definition (Modell)

Eine Interpretation I heisst **Modell** von φ wenn $I \models \varphi$.

Definition (erfüllbar etc.)

Eine Formel φ heisst

- ▶ **erfüllbar**, wenn es eine Interpretation I mit $I \models \varphi$ gibt.
- ▶ **unerfüllbar**, wenn φ nicht erfüllbar ist.
- ▶ **falsifizierbar**, wenn es eine Interpretation I mit $I \not\models \varphi$ gibt.
- ▶ **allgemeingültig** (= **gültig** = eine **Tautologie**), wenn $I \models \varphi$ für alle Interpretationen I gilt.

Definition (logische Äquivalenz)

Formeln φ und ψ heissen **logisch äquivalent** ($\varphi \equiv \psi$), wenn für alle Interpretation I gilt: $I \models \varphi$ gdw. $I \models \psi$.

Wahrheitstabelle

Wahrheitstabelle

Wie kann (Un-) Erfüllbarkeit, Falsifizierbarkeit, Allgemeingültigkeit automatisch für eine Formel bestimmt werden?

$\rightsquigarrow$ einfache Methode: **Wahrheitstabelle**

Beispiel: Ist $\varphi = ((P \vee H) \wedge \neg H) \rightarrow P$ allgemeingültig?

P	H	$P \vee H$	$((P \vee H) \wedge \neg H)$	$((P \vee H) \wedge \neg H) \rightarrow P$
F	F	F	F	T
F	T	T	F	T
T	F	T	T	T
T	T	T	F	T

$I \models \varphi$ für alle Interpretationen $I \rightsquigarrow \varphi$ ist allgemeingültig.

- ▶ Erfüllbarkeit, Falsifizierbarkeit, Unerfüllbarkeit?

Normalformen: Begriffe

Definition (Literal)

Wenn $P \in \Sigma$, dann heissen die Formeln P sowie $\neg P$ **Literale**.

P heisst **positives Literal**, $\neg P$ heisst **negatives Literal**.

Das **komplementäre Literal** zu P ist $\neg P$ und umgekehrt.

Wenn ℓ ein Literal ist, bezeichnet $\bar{\ell}$ das komplementäre Literal zu ℓ .

Definition (Klausel)

Eine Disjunktion von 0 oder mehr Literalen heisst **Klausel**.

Die **leere Klausel** $\perp$ wird auch als $\square$ geschrieben.

Klauseln aus nur einem Literal heissen **Einheitsklauseln**.

Definition (Monom)

Eine Konjunktion von 0 oder mehr Literalen heisst **Monom**.

Normalformen

Definition (Normalformen)

Formel φ ist in **konjunktiver Normalform** (KNF, Klauselform), wenn sie eine Konjunktion von 0 oder mehr Klauseln ist:

$$\varphi = \bigwedge_{i=1}^n \left(\bigvee_{j=1}^{m_i} \ell_{i,j} \right)$$

Formel φ ist in **disjunktiver Normalform** (DNF), wenn sie eine Disjunktion von 0 oder mehr Monomen ist:

$$\varphi = \bigvee_{i=1}^n \left(\bigwedge_{j=1}^{m_i} \ell_{i,j} \right)$$

Normalformen

Zu jeder Formel gibt es logisch äquivalente Formeln in KNF und in DNF.

Umwandlung in KNF

Wichtige Umformungsregeln für Umwandlung in KNF:

- ▶ $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi$ ($(\rightarrow)$ -Elimination)
- ▶ $\neg(\varphi \wedge \psi) \equiv \neg\varphi \vee \neg\psi$ (De Morgan)
- ▶ $\neg(\varphi \vee \psi) \equiv \neg\varphi \wedge \neg\psi$ (De Morgan)
- ▶ $\neg\neg\varphi \equiv \varphi$ (Doppelnegation)
- ▶ $(\varphi \wedge \psi) \vee \eta \equiv (\varphi \vee \eta) \wedge (\psi \vee \eta)$ (Distributivität)

Es gibt Formeln, bei denen jede äquivalente KNF- bzw. DNF-Formel exponentiell länger ist.

12.3 Logisches Schliessen

Logisches Schliessen: Intuition

Logisches Schliessen: Intuition

- ▶ Formeln repräsentieren im Allgemeinen nur eine unvollständige Beschreibung der Welt.
- ▶ Oftmals möchte man wissen, ob eine Formel aus einer (Menge von) anderen **logisch folgt**.
- ▶ Was bedeutet das?

Logisches Schliessen: Intuition

- ▶ **Beispiel:** $\varphi = (P \vee Q) \wedge (R \vee \neg P) \wedge S$
- ▶ S gilt in allen Modellen von φ .
Wie ist es für P , Q und R ?
- ↪ Betrachte alle Modelle für φ :

P	Q	R	S
F	T	F	T
F	T	T	T
T	F	T	T
T	T	T	T

Beobachtung

- ▶ In allen Modellen für φ gilt auch $Q \vee R$.
- ▶ Wir sagen: $Q \vee R$ **folgt logisch** aus φ .

Logisches Schliessen: Formal

Definition (Logische Folgerung)

Sei Φ eine Formelmenge. Eine Formel ψ **folgt logisch** aus Φ , in Symbolen $\Phi \models \psi$, wenn alle Modelle von Φ auch Modelle von ψ sind.

Anders gesagt muss für jede Interpretation I gelten: wenn $I \models \varphi$ für alle $\varphi \in \Phi$ gilt, dann muss auch $I \models \psi$ gelten.

Frage

Wie können wir automatisch berechnen, ob $\Phi \models \psi$ gilt?

- ▶ Eine Möglichkeit: Wahrheitstabelle. **Wie?**
- ▶ Geht es auch „besser“, d.h., (potentiell) ohne die komplette Wahrheitstabelle aufbauen zu müssen?

Logisches Schliessen: Deduktionssatz

Satz (Deduktionssatz)

Sei Φ eine endliche Menge von Formeln, ψ eine Formel. Dann gilt:

$$\Phi \models \psi \quad \text{gdw.} \quad \left(\bigwedge_{\varphi \in \Phi} \varphi \right) \rightarrow \psi \text{ ist eine Tautologie}$$

Beweis.

$\Phi \models \psi$

gdw. für alle Interpr. I : Wenn $I \models \varphi$ für alle $\varphi \in \Phi$, dann $I \models \psi$

gdw. für alle Interpr. I : Wenn $I \models \bigwedge_{\varphi \in \Phi} \varphi$, dann $I \models \psi$

gdw. für alle Interpr. I : $I \not\models \bigwedge_{\varphi \in \Phi} \varphi$ oder $I \models \psi$

gdw. für alle Interpr. I : $I \models (\bigwedge_{\varphi \in \Phi} \varphi) \rightarrow \psi$

gdw. $(\bigwedge_{\varphi \in \Phi} \varphi) \rightarrow \psi$ ist Tautologie □

Logisches Schliessen

Anwendung des Deduktionssatz

Logisches Schliessen kann auf Testen von Allgemeingültigkeit zurückgeführt werden.

Algorithmus

Frage: Gilt $\Phi \models \psi$?

- ① Prüfe, ob $(\bigwedge_{\varphi \in \Phi} \varphi) \rightarrow \psi$ eine Tautologie ist
- ② Wenn ja, dann $\Phi \models \psi$, ansonsten $\Phi \not\models \psi$

Im Folgenden: Wie kann Allgemeingültigkeit „effizient“ geprüft werden, d.h. ohne komplette Wahrheitstabellen zu bilden?

12.4 Resolution

Klauselmengen

Für den Rest des Kapitels:

- **Voraussetzung:** Formeln sind in konjunktiver Normalform
- Darstellung von Klauseln als **Menge C von Literalen**
- Darstellung von Formeln als **Menge Δ von Klauseln**

Beispiel

Sei $\varphi = (P \vee Q) \wedge \neg P$.

- φ ist in konjunktiver Normalform,
- φ besteht aus den Klauseln $(P \vee Q)$ und $\neg P$.
- Darstellung von φ als Menge von Mengen von Literalen:
 $\{\{P, Q\}, \{\neg P\}\}$

Unterscheide $\square$ (leere Klausel) vs. $\emptyset$ (leere Menge von Klauseln).

Resolution: Idee

Beobachtung

- Prüfen auf Allgemeingültigkeit kann auf Prüfen von Unerfüllbarkeit zurückgeführt werden.
- Formel φ allgemeingültig gdw. $\neg\varphi$ unerfüllbar.

Resolution: Idee

- Methode, um Unerfüllbarkeit einer Formel φ zu prüfen.
- **Idee:** Leite aus φ neue Formeln ab, die aus φ logisch folgen.
- Wenn leere Klausel $\square$ abgeleitet werden kann, dann ist φ unerfüllbar.

Die Resolutionsregel

$$\frac{C_1 \cup \{\ell\}, C_2 \cup \{\bar{\ell}\}}{C_1 \cup C_2}$$

- „Aus $C_1 \cup \{\ell\}$ und $C_2 \cup \{\bar{\ell}\}$ kann man $C_1 \cup C_2$ folgern.“
- $C_1 \cup C_2$ ist **Resolvent** der **Elternklauseln** $C_1 \cup \{\ell\}$ und $C_2 \cup \{\bar{\ell}\}$.
- Die Literale ℓ und $\bar{\ell}$ heissen **Resolutionsliterale**, die zugehörige Proposition **Resolutionsvariable**.
- Der Resolvent folgt logisch aus den Elternklauseln. (**Warum?**)

Beispiel

- Resolvent von $\{A, B, \neg C\}$ und $\{A, D, C\}$?
- Resolventen von $\{\neg A, B, \neg C\}$ und $\{A, D, C\}$?

Resolution: Ableitungen

Definition (Ableitung)

Notation: $R(\Delta) = \Delta \cup \{C \mid C \text{ ist Resolvent von 2 Klauseln in } \Delta\}$

Eine Klausel D kann aus Δ **abgeleitet** werden, in Symbolen $\Delta \vdash D$, wenn es eine Folge von Klauseln $C_1, \dots, C_n = D$ gibt, so dass für alle $i \in \{1, \dots, n\}$ gilt: $C_i \in R(\Delta \cup \{C_1, \dots, C_{i-1}\})$

Lemma (Korrektheit der Resolution)

Wenn $\Delta \vdash D$, dann $\Delta \models D$.

Gilt auch die Rückrichtung (**Vollständigkeit**)?

Resolution: Vollständigkeit?

Rückrichtung des Lemmas gilt nicht im Allgemeinen.

Beispiel:

- ▶ $\{\{A, B\}, \{\neg B, C\}\} \models \{A, B, C\}$, aber
- ▶ $\{\{A, B\}, \{\neg B, C\}\} \not\models \{A, B, C\}$

Aber: Rückrichtung gilt für Spezialfall der leeren Klausel $\square$.

Satz (Widerlegungsvollständigkeit der Resolution)

Δ ist unerfüllbar gdw. $\Delta \vdash \square$

Folgerung:

- ▶ Resolution ist ein vollständiges Beweisverfahren, um Unerfüllbarkeit zu testen.
- ▶ Resolution kann für logisches Schliessen eingesetzt werden, indem man auf einen Unerfüllbarkeitstest reduziert.

Resolution: Beispiel

Sei $\Phi = \{P \vee Q, \neg P\}$. Gilt $\Phi \models Q$?

Lösung

- ▶ Prüfe hierzu: Ist $((P \vee Q) \wedge \neg P) \rightarrow Q$ Tautologie?
- ▶ Äquivalent: Ist $((P \vee Q) \wedge \neg P) \wedge \neg Q$ unerfüllbar?
- ▶ Resultierende Klauselmengen Φ' : $\{\{P, Q\}, \{\neg P\}, \{\neg Q\}\}$
- ▶ Resolution von $\{P, Q\}$ mit $\{\neg P\}$ ergibt $\{Q\}$
- ▶ Resolution von $\{Q\}$ mit $\{\neg Q\}$ ergibt $\square$.
- ▶ Beobachtung: Leere Klausel ableitbar, also Φ^K unerfüllbar.
- ▶ Folgerung: $\Phi \models Q$

Resolution: Diskussion

- ▶ Resolution ist eine vollständige Beweismethode, um eine Formel auf Unerfüllbarkeit zu testen.
- ▶ Im schlechtesten Fall kann ein Resolutionsbeweis exponentiell lange dauern.
- ▶ In der Praxis ist eine Strategie hilfreich, welche Resolutionsschritte als nächstes angewandt werden.
- ▶ Im Folgekapitel lernen wir den DPLL-Algorithmus kennen, der als Kombination der Resolution mit Backtracking verstanden werden kann.

Zusammenfassung und Ausblick

Zusammenfassung

- ▶ Aussagenlogik bildet die Basis für allgemeine Repräsentation von Problemen und von Wissen
- ▶ In der Aussagenlogik kann man logisches Schliessen auf Testen von (Un-) Erfüllbarkeit reduzieren.
- ▶ Resolution ist eine vollständige Beweismethode, um Formeln auf Unerfüllbarkeit zu prüfen

Ausblick für Freitag

- ▶ Davis-Putnam-Logemann-Loveland-Methode (DPLL)