

BILDERARBEITUNG

FARBENMODELLE

Umrechnung RGB → HSI: Hue = Farbnuanze, Saturation = Sättigung, Intensity = Helligkeit

$$I = \text{avg}(R, G, B) = \frac{1}{3}(R + G + B)$$

$$S = 1 - \frac{\min(R, G, B)}{I} = 1 - \frac{3}{R + G + B} \min(R, G, B)$$

$$c = \arccos \frac{2R - G - B}{2\sqrt{(R - G)^2 + (R - B)(G - B)}}$$

$$H = \begin{cases} c & \text{falls } B < G \\ 360^\circ - c & \text{sonst} \end{cases}$$

- **Überschlagen:**
 - $R = 0^\circ = 0; G = 120^\circ \approx 85; B = 240^\circ \approx 170$
 - $I = 0 \rightarrow \text{Schwarz}$
 - $S = 0 \rightarrow \text{Grau}$
- Falls $R = G = B$, ist H undefiniert; falls $R = G = B = 0$, dann ist auch S undefiniert
- Helligkeit vom Farbwert getrennt → gegen Beleuchtungsänderungen unempfindlich

LOCHKAMERAMODELL

INTRINSISCHE KALIBRIERUNG

Kalibriermatrix	$K = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}$
Punktprojektion	$\begin{pmatrix} u \cdot w \\ v \cdot w \\ w \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = KP_c$
Inverse	$K^{-1} = \begin{pmatrix} 1/f_x & 0 & -c_x/f_x \\ 0 & 1/f_y & -c_y/f_y \\ 0 & 0 & 1 \end{pmatrix}$
Gerade in Kamera-KoSy	$g_c(w) = K^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} w = \begin{pmatrix} 1/f_x & 0 & -c_x/f_x \\ 0 & 1/f_y & -c_y/f_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} w$

EXTRINSISCHE KALIBRIERUNG

Matrix	$H = \begin{pmatrix} R & \mathbf{t} \\ 0 & 1 \end{pmatrix}$
Inverse	$H^{-1} = \begin{pmatrix} R^T & -R^T \mathbf{t} \\ 0 & 1 \end{pmatrix}$
Trafo Welt → Kamera	$\mathbf{x}_c = R\mathbf{x}_w + \mathbf{t}$
Projektionsmatrix	$P = (KR K\mathbf{t})$
Trafo Welt → Bild in homogenen Koordinaten	$\begin{pmatrix} u \cdot w \\ v \cdot w \\ w \end{pmatrix} = P \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = (KR K\mathbf{t}) \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$

ROTATIONEN

ROTATIONSMATRIZEN

2D		$P' = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = RP$
3D	Um X	$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$
	Um Y	$R_y(\theta) = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}$
	Um Z	$R_z(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Inverse		$R^{-1} = R^T$
Mitgedrehte Achsen (X→Y→Z)		$R_{X'Y'Z'}(\alpha, \beta, \gamma) = R_X(\alpha)R_Y(\beta)R_Z(\gamma)$
Raumfeste Achsen		$R_{ZYX}(\gamma, \beta, \alpha) = R_X(\alpha)R_Y(\beta)R_Z(\gamma)$
Nachteile		Hoch redundant, rechenaufwändig, Interpolation schwierig

QUATERNIONEN

Punkt → Quaternion	$P = (x, y, z) \rightarrow \mathbf{q}_P = (0, (x, y, z))$
Rotation	$\mathbf{q} = \left(\cos \frac{\theta}{2}, \mathbf{u} \sin \frac{\theta}{2} \right)$ mit \mathbf{u} = Einheitsvektor der gedrehten Achse
Konjugation	$\mathbf{q} = (q_w, \mathbf{q}_v) \rightarrow \bar{\mathbf{q}} = (q_w, -\mathbf{q}_v)$
Rotation mit Quaternionen	$\mathbf{q}'_P = \mathbf{q}\mathbf{q}_P\bar{\mathbf{q}}$
Norm	$N(\mathbf{q}) = \sqrt{q_w^2 + q_x^2 + q_y^2 + q_z^2}$
Multiplikative Inverse	$\mathbf{q}^{-1} = \frac{\bar{\mathbf{q}}}{N^2(\mathbf{q})}$
Nachteile	Keine Translation möglich

EPIPOLARGEOMETRIE

- **E-Geometrie** Zusammenhang zwischen zwei Kameras
- **Epipole** Schnittpunkte der Gerade durch beide Projektionszentren mit der jeweiligen Bildebene
- **E-Ebene** $\pi(X)$ = Ebene durch Epipole und den Punkt X
- **E-Linie** Schnitt der Epipolarebene mit der Bildebene; alle Punkte entlang dieser Linie werden auf ein Bildpunkt in der anderen Kamera projiziert
- Fundamentalmatrix F = mathematische Beschreibung der E-Geometrie. Eigenschaften: Matrix 3×3 , Rang 2 und für alle Korrespondenzen \mathbf{x}, \mathbf{x}' gilt $\mathbf{x}'^T F \mathbf{x} = 0$
- Nach **Rektifizierung** verlaufen alle E-Linien horizontal mit derselben v -Koordinate wie der Bildpunkt im anderen Kamerabild → Korrespondenzen nur noch horizontal

Epipole	$F\mathbf{e} = 0 \cap F^T\mathbf{e}' = 0$
Epipolarlinie	$l(\mathbf{x}') = F^T\mathbf{x}' \cap l'(\mathbf{x}) = F\mathbf{x}$
Essentialmatrix	$E = \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{pmatrix} R$ mit $(R \mathbf{t})$ = Transformation zw. Kamera 1 und 2
Fundamentalmatrix	$F = K'^{-T}EK^{-1}$
$F \rightarrow E$	$E = K'^T FK$

KONTRASTANPASSUNG

Affine Punktoperation	$I'(u, v) = \min(\max(\text{round}(a \cdot I(u, v)) + b, 0), q)$
Spreizung	$I'(u, v) = q \cdot \frac{I(u, v) - \min}{\max - \min}$
Histogramm	$H(x) := \#(u, v): I(u, v) = x, x \in [0, \dots, q - 1]$ $H_a(x) := \sum_{k=0}^x H(k)$ $H_n(x) := \text{round}\left(q \frac{H_a(x)}{H_a(q)}\right)$ $H_q(p) := \inf\{x \in \{0, \dots, q\}: H_a(x) \geq p \cdot H_a(q)\}$
H-Dehnung	$\min = H_q(p_{\min}), \max = H_q(p_{\max})$
H-Ausgleich	$I'(u, v) = H_n(I(u, v))$

HD braucht Quantile, HA nicht! HD erhöht den Kontrast der Graustufen zwischen gewählten Quantilen; HA erhöht den Kontrast in stark vertretenen Grauerbereichen, aber verringert ihn in schwach vertretenen. Kann auch zu Kontrastverminderung führen, z.B. (0 255) → (0 127)!

BILDFILTER

Die Transferfunktion $F(u, v)$ ist die Fourier-Transformierte der Filterfunktion $f(x, y)$

TIEFPASSFILTER

➔ Rauschunterdrückung, Glättung

Mittelwertfilter	Gauß
$\frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$	$\frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$ mit $\sigma = 0.85$

Größe der Matrix $n \times n$ für Gauß-Approximation:

$$n = \lfloor 2\sigma \rfloor \cdot 2 + 1$$

HOCHPASSFILTER

➔ Kantendetektion („X“ → vertikal, „Y“ → horizontal)

Achse	Prewitt	Sobel	Roberts	Laplace
X	$\begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}$
Y	$\begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	oder $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

SONSTIGES

- Laplacian of Gaussian (LoG): Laplacian ist sehr gegen Rauschen anfällig, deshalb zuerst mit Gauß glätten, dann mit Laplace die Kanten detektieren
- Canny-Kantendetektor
- Bandsperre: Setzt alle Werte im bestimmten Intervall auf null

SEGMENTIERUNG

- Schwellwertfilterung: Setzt alle Werte unter Schwellwert auf null, den Rest auf Max
- Farbe: Histogramme, morphologische Operationen

- Bewegung: Differenzbilder
- Region Growing: Fange mit einem Punkt an; finde ein Punkt im Region und ein Nachbar davon, so dass ihr Farbunterschied unter einem Schwellwert liegt; füge Nachbar dazu
- Kanten: Regressionsgeraden, Iterative Endpoint Fit, Hough-Transformationen
- Punktmerkmale: Korrespondenzproblem

MORPHOLOGISCHE OPERATOREN

Erosion	Nehme Pixel auf, nur falls alle unter der Maske gesetzt → Objekte werden verkleinert, dünne Linien verschwinden
Dilatation	Nehme Pixel auf, falls irgendein Pixel unter der Maske gesetzt → Objekte werden größer, Hohlräume werden geschlossen
Opening	Erosion, dann Dilatation → dünne Linien verschwinden
Closing	Dilatation, dann Erosion → Lücken werden geschlossen
Invertieren	Hilfstrick, um ohne einer anderen Operation auszukommen

HOUGH-TRANSFORMATION

Punkt	$r = x \cos \theta + y \sin \theta$
Gerade	$f(x) = -\frac{\cos \theta}{\sin \theta} x + \frac{r}{\sin \theta}$

KORRESPONDENZPROBLEM

Vergleiche zwei Bildausschnitte als $(2n + 1) \times (2n + 1)$ Matrizen.

SAD / SSD	Summiere alle absolute bzw. quadrierte Abweichungen → wird bei guter Übereinstimmung minimal
ZNCC (Kreuzkorrespondenz)	Normalisiert die Bitmaps bzgl. Beleuchtung → wird bei guter Übereinstimmung maximal (1)

SIGNALVERARBEITUNG

GRAPHISCHE FALTUNG

<https://www.youtube.com/watch?v=zoRJZDiPGds>

Flach auf Flach	Steigende Gerade
Flach auf Null	Fallende Gerade
Zack \geq auf Flach/Null	Negative Parabel
Zack \leq auf Flach/Null	Positive Parabel
Zack auf Zack	Häßlich

FOURIER-TRANSFORMATION

KONTINUIERLICHE TRANSFORMATION

Nur möglich, wenn $f(t)$ absolut integrierbar ist:

Transformation	$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$
Rücktransformation	$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{+i\omega t} d\omega$
Ab hier gilt: $f(t) \leftrightarrow F(\omega) \cap g(t) \leftrightarrow G(\omega)$	

Linearität	$af(t) + bg(t) \leftrightarrow aF(\omega) + bG(\omega)$
Zeitliche Verschiebung	$f(t + a) \leftrightarrow e^{i\omega a} F(\omega)$
Frequenzverschiebung	$f(t) \cdot e^{-iat} \leftrightarrow F(\omega + a)$
Skalierung	$f(at) \leftrightarrow \frac{1}{ a } F\left(\frac{\omega}{a}\right)$
Faltung	$f(t) * g(t) = \int f(\xi)g(t - \xi)d\xi$
Faltung und Fourier	$f(t) * g(t) \leftrightarrow F(\omega) \cdot G(\omega)$ $f(t) \cdot g(t) \leftrightarrow \frac{1}{2\pi} F(\omega) * G(\omega)$

DISKRETE TRANSFORMATION

Seien $f[k = 0, 1, \dots, N - 1]$ die N Messwerte und $F[\omega = 0, 1, \dots, N - 1]$ die Amplituden des Spektrums:

Transformation	$F[\omega] = \frac{1}{N} \sum_{k=0}^{N-1} f[k] e^{-i2\pi \frac{k}{N} \omega}$
Rücktransformation	$f[k] = \sum_{\omega=0}^{N-1} F[\omega] e^{+i2\pi \frac{\omega}{N} k}$
Frequenzauflösung	$\text{Frequenzauflösung} = \frac{f_{\text{Sampling}}}{N}$
Zeitliche Auflösung	$\text{Zeitauflösung} = \frac{\text{Frequenzauflösung}}{f_{\text{Sampling}}} = N^{-1}$

WICHTIGE FUNKTIONEN

Funktion	Fourier-Transformierte
$f(t) = e^{iat}$	$F(\omega) = \delta(\omega - \alpha)$
$f(t) = A \sin(2\pi \alpha t)$	$F(\omega) = \frac{A}{2} i\delta(\omega + \alpha) - \frac{A}{2} i\delta(\omega - \alpha)$
$f(t) = A \cos(2\pi \alpha t)$	$F(\omega) = \frac{A}{2} i\delta(\omega + \alpha) + \frac{A}{2} i\delta(\omega - \alpha)$
$f(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT)$	$F(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} \delta\left(\omega - \frac{n}{T}\right)$

NYQUIST-SHANNON ABTASTTHEOREM

- Abtastung des Signals $f(t)$ mit der Periode T entspricht der Multiplikation davon mit der unendlichen Impulsfolge $\sum_{n=-\infty}^{\infty} \delta(t - nT)$.
- Problem: **Aliasing**. Bei Sampling mit zu geringer Abtastrate überlagern sich zu hohe Frequenzen im Spektrum mit sich selbst und lassen die Originalfrequenz nicht mehr durch die inverse Fourier-Transformation rekonstruieren.
- Lösung: Die Samplingfrequenz muss echt größer als die zweifache im Signal vorkommende Frequenz (Cutoff-Frequency) sein:

$$\Delta x < \frac{1}{2\omega} \Leftrightarrow f_{\text{Sampling}} > 2f_{\text{Cutoff}}$$

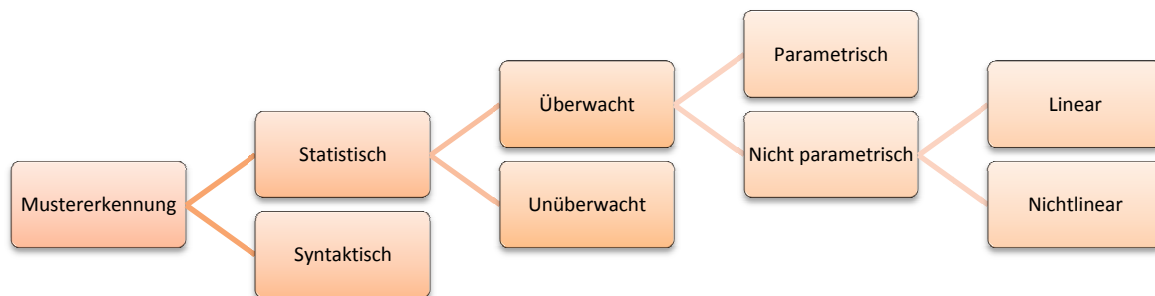
KLASSIFIKATION

- Fluch der Dimensionalität: Unreflektiertes Hinzufügen von Objektfeatures führt zum schlechteren Performance, weil zu viele Parameter (Dimensionen), zu wenig Daten.

KLASSIFIKATOR-ARTEN

- Syntaktisch = eine Syntaxmodell wird eingegeben; statistisch = reine Daten
- Überwacht = Daten sind davor gelabelt; unüberwacht = Features und Klassen müssen automatisch extrahiert werden
- Parametrisch = eine unterliegende Wahrscheinlichkeitsverteilung wird angenommen (z.B. Gauß) und über deren Parameter wird optimiert; nicht parametrisch = keine Verteilung angenommen

Verfahren	Nomenklatur			
Clustering	Statistisch	Unüberwacht		
Deep Learning	Statistisch	Unüberwacht		
Mixture Densities	Statistisch	Unüberwacht		
Gaussian Classifier	Statistisch	Überwacht	Parametrisch	
K-Nearest Neighbors	Statistisch	Überwacht	Nicht param.	
Parzen Windows	Statistisch	Überwacht	Nicht param.	
Linear Discriminant	Statistisch	Überwacht	Nicht param.	Linear
Perzeptron	Statistisch	Überwacht	Nicht param.	Linear
Template Matching	Statistisch	Überwacht	Nicht param.	Linear
MLP + ANN	Statistisch	Überwacht	Nicht param.	Nichtlinear



BAYES-KLASSIFIKATOR

Statistisch, überwacht, parametrisch

Bayes-Regel	$p(w x) = \frac{p(x w) \cdot P(w)}{P(x)}$ <p>mit x = Messung und w = Klasse</p>
A-Priori Klassen-wahrscheinlichkeit	$P(w) = \frac{\#(x_{Training} \in w)}{\#(x_{Training})}$
Bedingte Messungs-wahrscheinlichkeit (Gauß)	$p(x w) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ <p>mit μ = Mittelwert aller Trainingswerte von der Klasse und σ^2 = Varianz der Trainingswerte von der Klasse</p>
Maximum-Likelihood-Estimation	$w^*(x) = \operatorname{argmax}_{w} (p(w x)) = \operatorname{argmax}_{w} (p(x w) \cdot P(w))$
Fehlerwahrscheinlichkeit bei Schwellwert θ	$p_{\theta}(\text{Fehler}) = \int_{-\infty}^{\theta} p(w_2 x) dx + \int_{\theta}^{\infty} p(w_1 x) dx$
Opt. Schwellwert für 2 Klassen /Entscheidungsgrenze	$\theta_{opt}: p(w_1 \theta_{opt}) = p(w_2 \theta_{opt})$ $\ln(P(w_1) \cdot p(\theta_{opt} w_1)) = \ln(P(w_2) \cdot p(\theta_{opt} w_2))$

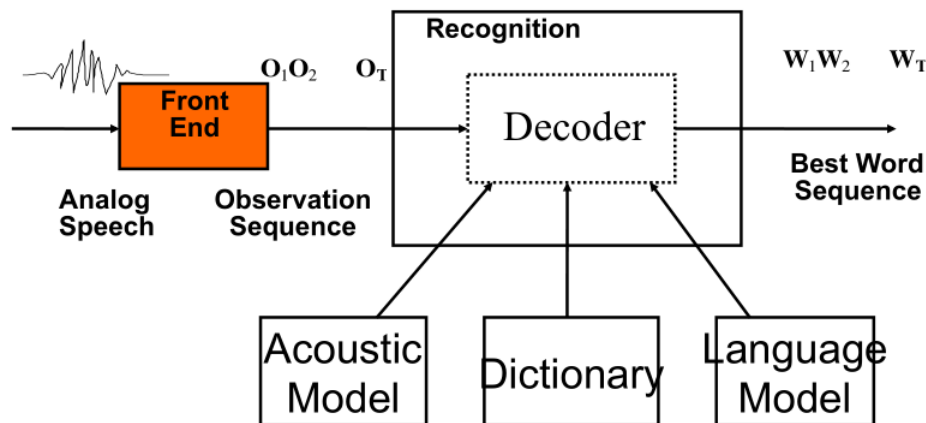
PERZEPTRONEN

Statistisch, überwacht, nicht parametrisch, linear

- XOR lässt sich erst mit einer versteckten Schicht mit mind. 2 Perzeptronen lösen

Trenngerade von einem Perzeptron mit 2 Eingänge	$w_1x + w_2y + b = 0$ $f(x) = -\frac{w_1}{w_2} - \frac{b}{w_2}$
---	---

SPRACHVERARBEITUNG



- Fundamentalformel der Spracherkennung:

$$W^* = \operatorname{argmax}_{W \in \mathcal{W}} P(W|X) = \operatorname{argmax}_{W \in \mathcal{W}} \frac{P(X|W) \cdot P(W)}{P(X)} = \operatorname{argmax}_{W \in \mathcal{W}} P(X|W)$$
- Vorverarbeitung: Anti-Aliasing Filter, Analog-Digital-Umwandlung, FFT, usw.
- Decoder: $\operatorname{argmax}_{W \in \mathcal{W}}$
 - Akustisches Modell $p(X|W)$ → HMM, Gaussian Mixture, TTNN
 - Wörterbuch \mathcal{W}
 - Sprachmodell $P(W)$ → grammatikbasiert oder n-Gramme
- Performancemessung: Alignment-Suche zwischen Hypothese und der Referenz:

Word Error Rate	$WER = \frac{\#Ins + \#Del + \#Sub}{N}$ <p>mit N = Anzahl Wörter im Referenzsatz</p>
Normalisierte Logprob eines Satzes W_a	$H(W_a) = -\frac{1}{ W_a } \sum_{i=1}^{ W_a } \log_2 P(w_i \Psi(w_1, \dots, w_{i-1}))$
Perplexität	$PPL(W_a) = 2^{H(W_a)}$

HIDDEN-MARKOV-MODELLE

- In der Spracherkennung: Zustände sind Phoneme, Ausgaben sind Beobachtungen
- Aufpassen: Ob das Signal im Modell vor oder nach dem Zustandsübergang emittiert wird
- Drei Hauptprobleme:
 - Evaluation mit dem Forward-Algorithmus: Berechne die Wahrscheinlichkeit einer Ausgabesequenz über alle mögliche Zustandssequenzen
 - Decoding mit dem Viterbi-Algorithmus: Finde die wahrscheinlichste Zustandssequenz zu einer Ausgabesequenz (wie Forward, aber Max statt Summe)
 - Training mit dem Forward-Backward-Algorithmus: Optimierung der Parameter

PLANUNG

LOGIK

- Normalformen aus der Wahrheitstabelle:
 - DNF: Überall, wo der Ausdruck zu „wahr“ evaluiert, die Literale verunden und als Klausel an die Disjunktion dranhängen
 - KNF: Überall, wo der Ausdruck zu „falsch“ evaluiert, die Literale negieren und verodern, und anschließend an die Konjunktion dranhängen
- Eine Horn-Klausel ist eine Disjunktion mit höchstens einem positiven Literal. Arten:
 - Definition 1+ negierte Literale, 1 positives Literal
 - Integritätseinschränkung 1+ negierte Literale, kein positives Literal
 - Fakt/Axiom keine negierte Literale, 1 positives Literal
- Resolutions-Algorithmus überprüft, ob eine Klausel aus der bestehenden Klauselmenge abgeleitet werden kann. Dazu wird die Zielklausel negiert hinzugefügt und anschließend mit der Resolutionsregel versucht, ein Widerspruchsbeweis zu finden (eine leere Klausel abzuleiten).
- DPLL-Algorithmus überprüft, ob eine Klauselmenge erfüllbar ist. Wichtige Begriffe:
 - **Einheitsklausel** enthält nur ein Literal → muss wahr sein
 - **Reine Variable** kommt in der ganzen Formel nur positive oder nur negiert vor

STRIPS

- „Geschlossene Welt“ Annahme: Variablen die nicht explizit wahr sind, sind werden als falsch angenommen.
- Zielzustandsdefinition: nur Konjunktion positiver Literale!
- Aktion = Name + Parameter (Variablen) + Vorbedingungen + Effekte (V. und E. können Variablen aus der Parameterliste enthalten)
- Effekte sind die einzige Stelle, an der negative Literale explizit vorkommen; nicht in der Effekt vorkommende Literale werden einfach nicht verändert
- **ADL**: Erweiterung von STRIPS
 - „Offene Welt“ Annahme: Nicht angegebene Literale gelten als unbekannt
 - Zielzustand kann Negationen und Disjunktionen enthalten
 - Gleichheitsoperator
 - Implizite Typisierung bei der Aktionsdefinition

PLANUNGSGRAPH

- Ergebnisse sind gute Heuristiken für A* u. ä.
- Zeigt ein mögliches Verlauf der Zustände und Aktionen an bzw. ab wann welche Variablen wahr werden können
- Mutex-Links: Verbinden sich gegenseitig ausschließende Teilzustände, sowie Aktionen, die nicht zusammen oder nacheinander ausgeführt werden können

UMWELTREPRÄSENTATION

- Polygonzerlegung: Vertikale Linien durch die Eckpunkte
- Quadrees: Rekursive in Quadrate zerteilen
- Voronoi-Diagramme: Region enthält alle Punkte, die zum Hindernis am nächsten sind
- Potentialfelder: Gradient Descent → gegen lokale Minima anfällig
- Sichtgraphen: Von jeder Ecke jede andere sichtbare Ecke verlinken