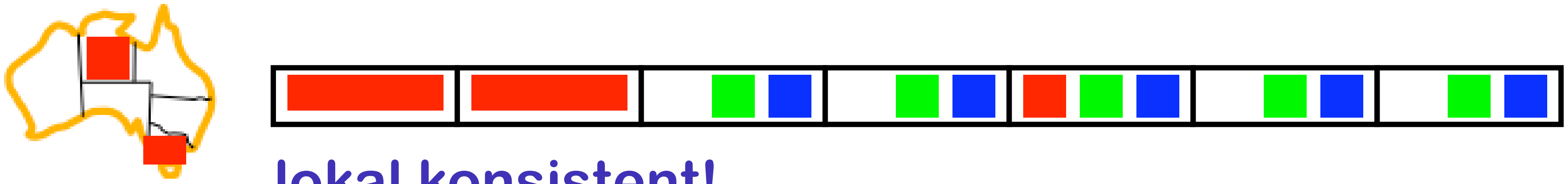
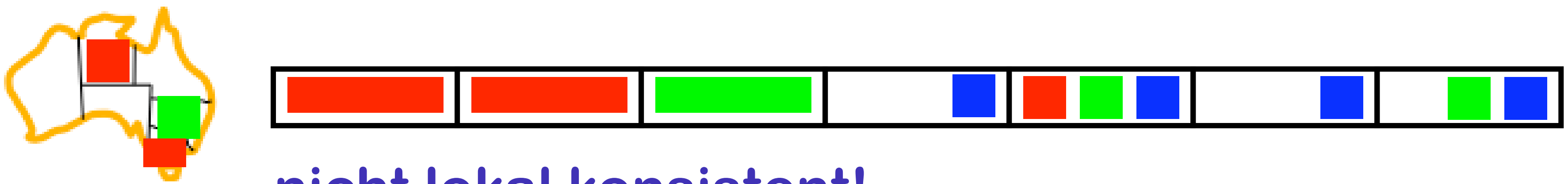
Beispiele im Färbeproblem

lokal konsistent!



lokal konsistent!



nicht lokal konsistent!

Q und SA können nicht beide blau sein

⇒ Kante zwischen Q und SA nicht konsistent

Der Algorithmus AC-3

… für Binärconstraints (jeder Constraint entspricht e. Kante, *arc*)

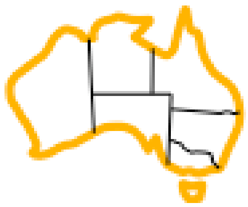
function AC3(csp) **returns** the CSP, possibly with reduced domains
  local variables: $queue$, a queue of arcs, initially all the arcs in csp

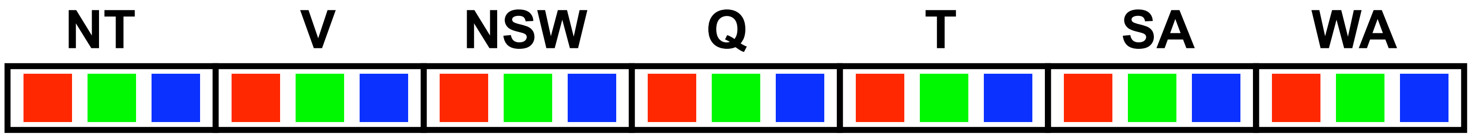
  loop while $queue$ is not empty **do**
    $(X_i,\ X_j) \leftarrow$ REMOVE-FRONT($queue$)
    if REMOVE-INCONSISTENT($X_i,\ X_j$) **then**
      for each X_k **in** NEIGHBORS[X_i] **do**
        add ($X_k,\ X_i$) to $queue$

---

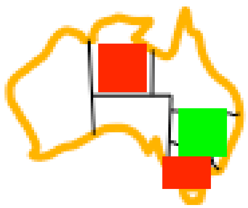
function REMOVE-INCONSISTENT($X_i,\ X_j$) **returns** true iff we remove a value
  $removed \leftarrow false$
  loop for each x **in** DOMAIN[X_i] **do**
    if no y in DOMAIN[X_j] allows (x,y) to fulfill the constraint between X_i, X_j
      then delete x from DOMAIN[X_i]; $removed \leftarrow true$
  return $removed$

Beispiel: AC-3 in Australien





... bleibt wie es ist!



... ergibt:

Eigenschaften des AC-3-Algorithmus

☺ Speicher: binäres Constraintnetz mit $O(n^2d^2)$
(n Variable, max. d Werte)

😐 Zeitbedarf für binäre Constraints: $O(n^2d^3)$
(für $O(n^2)$ Constraints packe max. d Mal d^2 Wertetupel an)

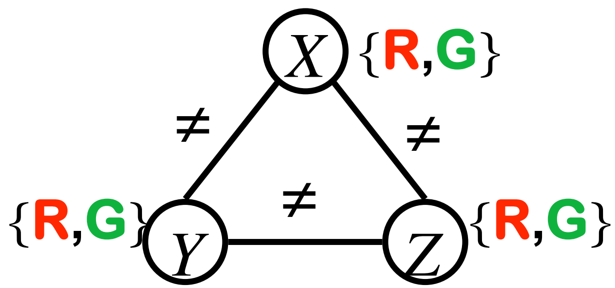
- Zeitbedarf verbesserbar auf $O(n^2d^2)$
(Führe Buch, welche Löschungen Nachbarvariablenwerte beeinträchtigen)
- Reduktion der Komplexität bei Zyklenfreiheit (z.B. Baum) und unzusammenhängenden Teilnetzen
(Australienbeispiel: Tasmanien unabhängig vom Rest färbbar)
- AC-3 terminiert unabhängig von der Sortierung in *queue* mit lokal konsistenter Belegung oder mit Ø für alle Variablen eines maximalen zusammenhängenden Teilnetzes

Lokale Konsistenz und Lösbarkeit

Polynomielle Laufzeit lässt vermuten, dass AC-3 im Allg. Suchprobleme nicht lösen kann (das geht nur exponenziell!):

Nicht alle Belegungen im lokal konsistenten Netz sind Lösungen!
Z.B.: Australien-Färbeproblem mit {R,G,B} f. a. Variablen ist lokal konsistent

Es gibt lokal konsistente Netze, die gar keine Lösung haben!
Beispiel:



Schärfere Konsistenzbegriffe erweitern den „Horizont“ der lokalen Konsistenz: k-Konsistenz, strenge k-Konsistenz

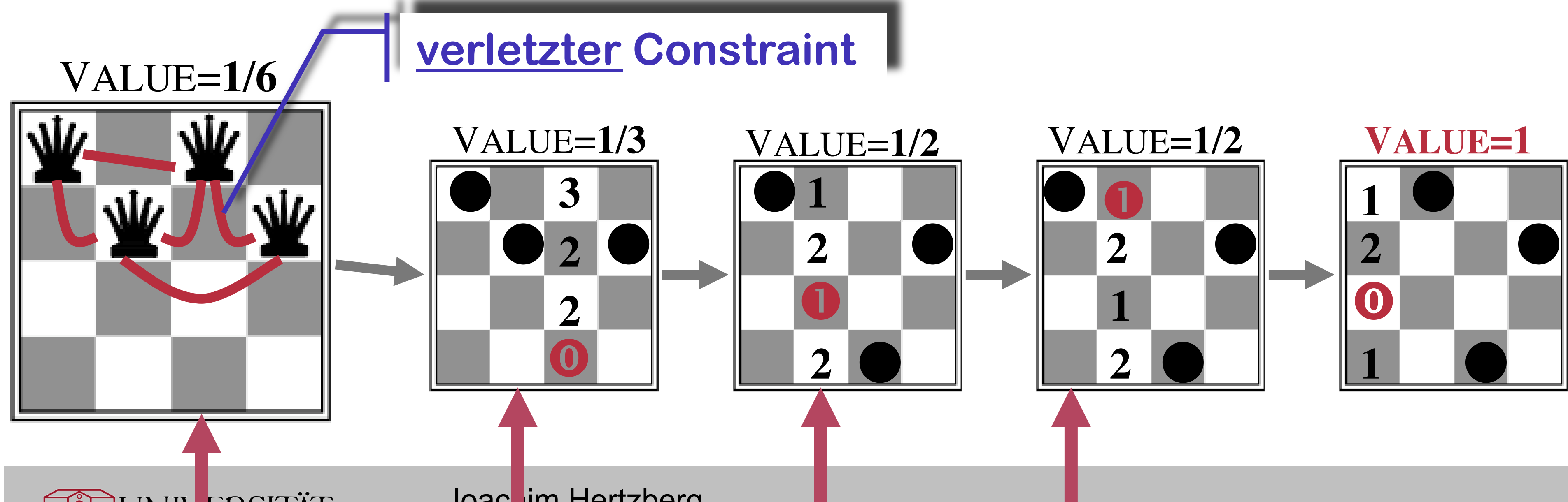
Lokale Suche für CSPs: Min-Conflicts

Für CSPs wie für Suche gibt es systematische und lokale Verfahren

Verwende Constraintnetz in d. <u>Zustandsbewertungsfunktion</u> bei lokaler Suche: VALUE(n) = 1/(1+#Constraintverletzungen)

Vergleiche n-Damen-Beispiel für Bergsteige-Suche

Verändere aktuellen Knoten so, dass für zufällig gewählte Variable die Anzahl d. Constraintverletzungen minimiert wird.



Min-Conflicts

function MIN-CONFLICTS(*csp*,*max_steps*) **returns** a solution or failure
inputs: *csp*, a constraint satisfaction problem
max_steps, the number of steps allowed before giving up

current ← an initial complete assignment for *csp*
for *i*=1 **to** *max_steps* **do**
  if *current* is a solution for *csp* **then return** *current*
  var ← a randomly chosen, conflicted variable from VARIABLES[*csp*]
  value ← a value *v* for *var* that minimizes CONFLICTS(*var*, *v*, *current*, *csp*)
  set *var*=*value* in *current*
return failure

- Vergleiche Algorithmus Bergsteigen, möglicherweise mit gleich bewertetem Nachfolger; vergleiche Random Walk
- Hier Lösungstest & Bewertung in Termini von Constraints

Zusammenfassung Constraints

- Constraintnetze repräsentieren Variablen, zwischen denen Relationen gelten
- Eine Lösung besteht in gültiger Belegung aller Variablen
- Constraint-Algorithmen sollen eine effiziente Suche nach einer solchen Belegung realisieren
- Varianten lokaler Konsistenz stellen Zwischenschritte zu e. Lösung dar, die auch anderen Suchverfahren dienen können
- Wie bei der allgemeinen Suche gibt es systematische und lokale Verfahren, eine Lösung zu suchen
- Praktisch werden Constraint-Verfahren oft in Zuordnungs-, Konfigurations- und Scheduling-Problemen eingesetzt
  ↪ *Constraint Logic Programming*, CLP ↪ SWI-Prolog!

4. Schließen unter Unscherheit

1. Was ist KI?
2. Logik und Inferenz
3. Suche als Problemlöseverfahren
4. **Schließen unter Unsicherheit**
5. Maschinelle
6. Ausblick: „

4.1 Unsicherheit und Wahrscheinlichkeit
4.2 Bayes-Netze

4.1 Unsicherheit und Wahrscheinlichkeit

- Die Effekte von Handlungen in der Welt unterliegen vielerlei Unsicherheiten, die der Agent beim Planen bedenken sollte.
- Beispiel: Aktion $A(n)$ bringe Agenten in n Minuten zum Flughafen. Unsicherheiten zur Planungszeit:
  - **Begrenzte Wahrnehmbarkeit** der Welt (Straßenzustand?)
  - **„Sensorrauschen“** (Verlässlichkeit von Staumeldungen?)
  - **Ausnahmefälle** von Aktionen (Platter? Benzin alle? Auto weg?)
  - **Komplexität** des zu modellierenden Weltausschnitts
- Logik-Ansätze zwingen dazu, sie „auszumodellieren“, wenn man sie berücksichtigen können will (s. 2.5)
- Das geht praktisch oft nicht, aber

Unsicherheit soll repräsentierbar sein!

Verwendung unsicherer Info. beim Planen

Unterschiedliche Aktionen eines Agenten erzeugen mit unterschiedlichen W'keiten Effekte (Sicht zur Planungszeit).

Beispiel Flughafenfahrt-Aktion Folie 195:

$P(A(25)$ bringt mich rechtzeitig zum Flug | $\ldots) = 0{,}04$
$P(A(90)$ bringt mich rechtzeitig zum Flug | $\ldots) = 0{,}8$
$P(A(120)$ bringt mich rechtzeitig zum Flug | $\ldots) = 0{,}95$
$P(A(1440)$ bringt mich rechtzeitig zum Flug | $\ldots) = 0{,}9999999999999999$

Und nun?

Handle so, dass Du den maximalen erwarteten Nutzen erzielst!
(s. Kapitel 1 (Folie 17): Rationaler Agent)

Entscheidung ≈ Effektw'keit x Nutzen

Verfügbare Formalismen

- Allerlei ad-hoc-Ansätze („unscharfe“ Regeln, …):
  keine ordentliche Theoriebasis
- **Normalfallschließen**: s. Abschnitt 2.5:
  oft zu genau, zu feinteilig, zu mächtig, daher zu aufwändig
- ***Fuzzy logic***: Modellierung von „gradueller Wahrheit“:
  $Groß(x)$ ist absolut wahr $(1{,}0)$ für $Peter$ (der 2,01m groß ist), mittelwahr $(0{,}4)$ für $Fritz$ (1,75m), unwahr $(0{,}05)$ für $Karl$ (1,52m)
  anderes Weltbild als Logik, daher andere Art Modellierung mit anders interpretierbaren Ergebnissen;
  in vielen Fällen erfolgreich angewendet
- **Wahrscheinlichkeitstheoretische** Ansätze:
  … um die es im Folgenden geht

Wahrscheinlichkeit (Sicht der KI)

… drückt (subjektive) Erwartung über die relative Häufigkeit von Fakten oder Zusammenhängen aus

… beruht auf dem Logik-Weltbild, dass Dinge „objektiv“ „eigentlich“ genau wahr oder falsch sind

… modelliert unterschiedliche Formen von Nichtwissen (theoretisches, praktisches, pragmatisches) in ein und demselben numerischen Maß („Wahrscheinlichkeit“)

Beispiel: **Korrelation** zwischen Loch im Zahn und Zahnschmerzen

Weder ist korrekt $\forall p.\, Symptom(p, Zahnschmerz) \Rightarrow Krankheit(p, Zahnloch)$

noch $\forall p.\, Krankheit(p, Zahnloch) \Rightarrow Symptom(p, Zahnschmerz)$

Ob Patient Loch im Zahn und ob er Zahnschmerzen hat, ist feststellbar. Genaue Zusammenhänge können unbekannt oder zu komplex sein.

W ‘theorie Grundbegriffe: Ereignis, Variable

Ergebnisraum Ω: Mögliche Ergebnisse eines Zufallsexperiments

z.B.: {„1“, „2“, „3“, „4“, „5“, „6“} beim Würfeln

{[„1“,*sonnig*], [„1“,*regnet*], [„2“,*sonnig*], …, [„6“,*regnet*]}

Ereignis:

Teilmenge v. Ω. z.B.: {„2“, „4“, „6“} beim Würfeln („gerade“)

Elementarereignis:

$\omega\in\Omega$. (eigentlich: einelementiges Ereignis) z.B.: [„1“,*sonnig*]

Zufallsvariable: Funktion von Elementarereignissen $\omega\in\Omega$ in einen Wertebereich $\mathcal{D}$. (Notation: beginnt mit Großbuchstabe)

$\mathcal{D}=\{\text{true}, \text{false}\}$: **Boolesche** Zufallsvariable, z.B.: *Zahnloch*

$\mathcal{D}$ diskret: **Diskrete** ZV, z.B.: *Augenzahl* (auf einem Würfel)

$\mathcal{D}$ stetig (hier: $\Re$): **kontinuierliche** ZV, z.B.: *Position* (Roboter im 2D)

W 'theorie Grundbegriffe: Wahrscheinlichkeitsraum

Wahrscheinlichkeitsraum (Ω, P), $\Omega \neq \emptyset$: P ist eine so auf Ereignissen aus Ω definierte reellwertige Funktion („Wahrscheinlichkeit"), dass die **Kolmogorow-Axiome** gelten:

1. $0 \leq P(e) \leq 1$ für alle $e \subseteq \Omega$
2. $P(\Omega) = 1$
3. Für $d, e \subseteq \Omega$, sodass $d \cap e = \emptyset$: $P(d \cup e) = P(d) + P(e)$

z.B.: $P(\text{„}gerade\text{“}) = P(\text{„}2\text{“}) + P(\text{„}4\text{“}) + P(\text{„}6\text{“}) = 1/6+1/6+1/6=1/2$

(W'raum präziser (Ω, Σ, P) mit Ereignisraum Σ – hier verkürzt dargestellt)

Ereignisse a,b sind **unabhängig**, gdw: $P(a \cap b) = P(a) \cdot P(b)$

W'theorie Grundbegriffe: W'verteilungen, -dichten

P induziert Funktionen auf den Zufallsvariablen über (Ω,P):

Wahrscheinlichkeitsverteilung $\mathbf{P}$ einer diskreten Zufallsvariablen X: Liste aus W'keiten der möglichen Werte für X, wobei $\sum_{x\in \mathrm{Range}(X)\subseteq\Omega}\mathbf{P}(x)=1$

z.B.: $\mathbf{P}(Augenzahl) = \langle P(„1“), \ldots, P(„6“)\rangle = \langle\, 1/6, 1/6, 1/6, 1/6, 1/6, 1/6 \,\rangle$

Notation: $\mathbf{P}(Augenzahl = „1“) = 1/6$

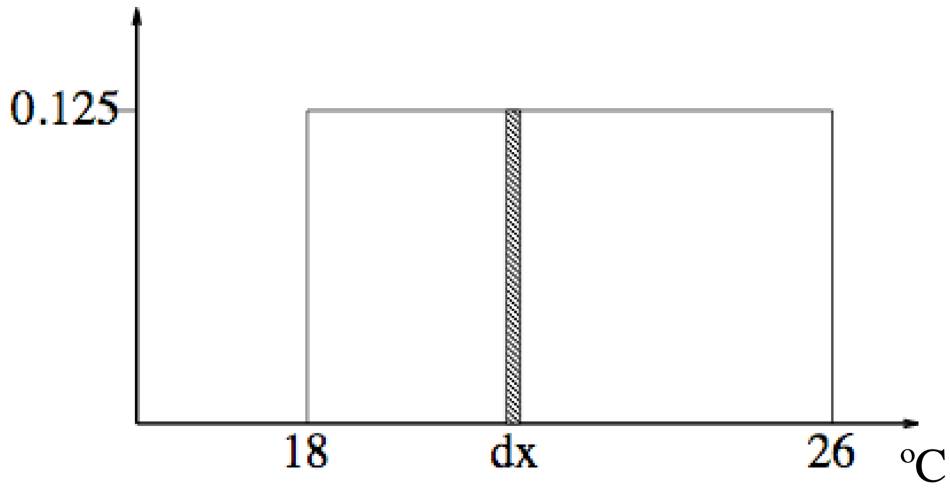
Wahrscheinlichkeitsdichte $\mathbf{P}$ einer kontinuierlichen Zufallsvariablen X: Fkt. vom Typ $\mathrm{Range}(X)\rightarrow[0,1]$, wobei

$$\int_{x\in \mathrm{Range}(X)\subseteq\Omega}\mathbf{P}(x)\mathrm{d}x = 1$$

und $\mathbf{P}(a \leq x \leq b) = \int_a^b \mathbf{P}(x)\,\mathrm{d}\,x$

z.B.: ...

Beispiel für W ‘dichte: Gleichverteilung



P(*Temperatur*) = *U*[18,26](*x*) =
Gleichverteilung auf [18,26]

$$\int_{-\infty}^{\infty} U[18{,}26](x)\mathrm{d}x = \int_{18}^{26} U[18{,}26](x)\mathrm{d}x = 1$$

$$\mathbf{P}(20 \leq Temperatur \leq 22) = \int_{20}^{22} U[18{,}26](x)dx = 0{,}25 = P([20{,}22])$$

$$P(20{,}5) = 0$$

$$\mathbf{P}(Temperatur{=}20{,}5) := \lim_{dx\to 0}\mathbf{P}(20{,}5 \leq X \leq 20{,}5+\mathrm{d}x)/\mathrm{d}x = 0{,}125/^{\circ}\mathrm{C}$$

W'theorie Grundbegriffe: Unbedingte W'keiten

W'keit P „an sich", ohne weitere Information:

Unbedingte Wahrscheinlichkeit (**a priori** W'keit)

z.B.: $P(regnet) = 0{,}3$

Verteilung für unbedingte W'keiten wie beschrieben

z.B.: $\mathbf{P}(OSWetter) = \langle P(sonnig), P(wolkig), P(regnet), P(frostig) \rangle$
$= \langle 0{,}25, 0{,}4, 0{,}3, 0{,}05 \rangle$

Gemeinsame Verteilung von n ZVen entspricht n-dim. Tabelle:

Notation: $\mathbf{P}(ZV_1 \cap ZV_2)$
oder $\mathbf{P}(ZV_1, ZV_2)$
Beispiel:
$\mathbf{P}(OSWetter, Zahnloch)$

OSWetter= *Zahnloch=*	*sonnig*	*wolkig*	*regnet*	*frostig*
false	0,225	0,36	0,27	0,045
true	0,025	0,04	0,03	0,005

Vollständige gem. Verteilg.: W'keit <u>aller</u> Elementarereignisse
(hat $O(m^n)$ Einzelw'keiten für n ZVen mit max. m Werten)

W'theorie Grundbegriffe: Bedingte W'keiten

W'keit P bei gegebener Information:

Bedingte Wahrscheinlichkeit (**a posteriori** W'keit)

z.B.: $P(regnet \mid luftdruck\_hoch) = 0{,}1$

Definiert (falls $P(b)>0$) als

($P(a|b)=0$, falls $P(b)=0$)

$$P(a \mid b) := \frac{P(a \cap b)}{P(b)}$$

… oder als („Produktregel“): $P(a \cap b) = P(a|b)P(b) = P(b|a)P(a)$

z.B.: $P(sonnig|Zahnloch{=}false) = 0{,}225/0{,}9 = 0{,}25$

entspricht $P(sonnig)$, weil unabhängig! ($P(a \cap b) = P(a)P(b)$)

Entsprechend für W'verteilungen:

$$\mathbf{P}(OSWetter, Zahnloch) = \mathbf{P}(OSWetter|Zahnloch)\,\mathbf{P}(Zahnloch)$$