Evolutionäre Algorithmen

Vorlesung 5

Schema-Theorem

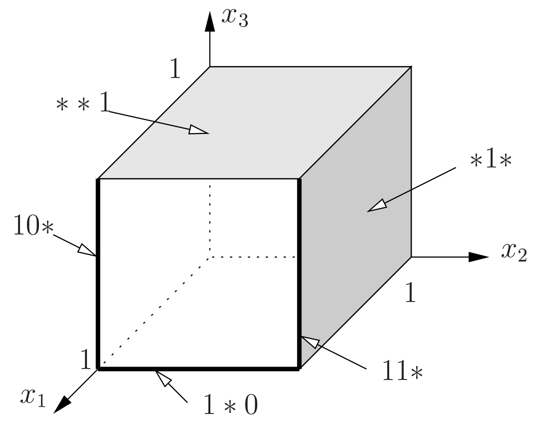
Formae

Price-Theorem

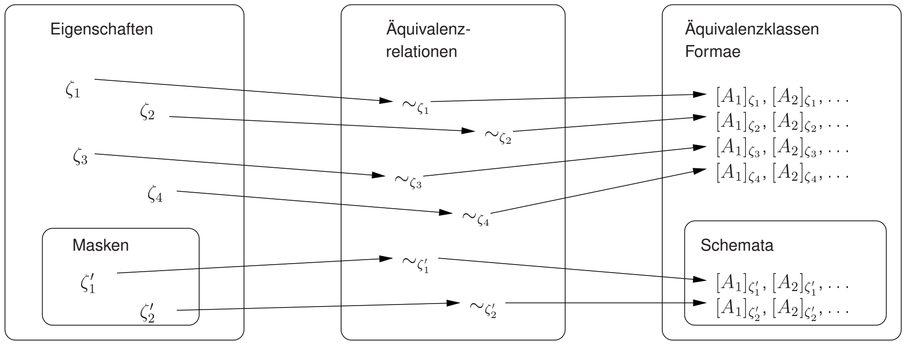
Suche in Hyperebenen

- ▷ Lösungskandidaten lassen sich durch unterschiedliche Eigenschaften charakterisieren und in unterschiedliche Klassen einteilen
- ▷ jede Population macht über die Gütewerte eine Aussage bezüglich des Nutzens der Eigenschaften
- ▷ Selektion favorisiert „gute“ Eigenschaften, die dann rekombiniert werden

Beispiel



Begriffe und Zusammenhänge



Merkmale und Formae

- Menge M von Merkmalen an Individuen
- Merkmal $\zeta \in M$ kann unterschiedliche
  Ausprägungen annehmen
  $\Rightarrow$ induziert eine Äquivalenzrelation $\sim_\zeta$
- Äquivalenzklassen $[A]_\zeta \;=\; \{B | A \sim_\zeta B\}$
  (heißen Formae)

Eigenschaften

- Anzahl der Äquivalenzklassen bzgl. eines Merkmals:
  Genauigkeit
- Formae ν und ν' sind *miteinander verträglich*
  $\nu \bowtie \nu' \;:\Leftrightarrow\; \nu \cap \nu' \neq \emptyset$
- Beispiel?

Positionsbasierte Merkmale

- Eigenschaften über Positionen im Genotyp definiert
- Maske $\tilde{\zeta} \in \mathcal{P}^{\{1,\ldots,l\}}$ bestimmt relevante Positionen
- Äquivalenzrelation $A \sim_{\tilde{\zeta}} B \;:\Leftrightarrow\; \forall_{i \in \tilde{\zeta}}\, A_i = B_i$
- Ordnung einer Maske $o(\tilde{\zeta})$:
  Anzahl der definierten Positionen
- definierende Länge einer Maske $\delta(\tilde{\zeta})$:
  $\delta(\tilde{\zeta}) \;:=\; \max\{|i-j| \mid i,j \in \tilde{\zeta}\}$

Schemata

- $\mathcal{G} = \{0,1\}^l$
- Beispiel: Masken $\tilde{\zeta} = \{2,4,5,6\}$ und $\tilde{\zeta} = \{1,5\}$ auf
  $\mathcal{G} = \{0,1\}^6$: Formae zum Individuum 100010?
- Schema-Schreibweise:
  $H \;=\; (H_1, \ldots, H_l) \;\in\; (\mathcal{G} \cup \{*\})^l$
  mit $H_i \;=\; \begin{cases} A_i, & \text{falls } i \in \tilde{\zeta} \\ , & \text{sonst} \end{cases}$
- Instanzen eines Schemas $I(H)$

Eine Generation eines EA

- Population $P^{(t)} = \left\langle A^{(t,i)} \right\rangle_{1\leq i\leq n}$, Forma v
- Elternselektion $Sel^{\xi,dec,f} : \mathcal{G}^n \to \mathcal{G}^n$,

  $E_{Sel}(v, P^{(t)}) \;:=\; \sum_{\xi\in\Xi} |v \cap Sel^{\xi,dec,f}|$
- $R^{\xi} : \mathcal{G}^2 \to \mathcal{G}^2$,

  $P_{R,v} \;=\; \Pr_{\xi\in\Xi, B\in\mathcal{G}} \left[R^{\xi}(A,B) \notin v \mid A \in v\right]$
- $M^{\xi} : \mathcal{G} \to \mathcal{G}, \;\; P_{M,v} \;=\; \Pr_{\xi\in\Xi} \left[M^{\xi}(A) \notin v \mid A \in v\right]$
- Wirkung auf Häufigkeit der Vertreter von v:

  $E(v, t+1) \;\geq\; E_{Sel}(v, P^{(t)})\,(1 - P_{M,v} - P_{R,v})$

Schema-Theorem

- mit fitnessproportionaler Selektion,

  bitflipping Mutation mit Inversionswahrscheinlichkeit p_M,

  1-Punkt-Crossover als Rekombination
- Dann gilt für $E(H,t) = |I(H) \cap P^{(t)}|$:

  $$E(H, t+1) \;\geq\; E(H,t)\frac{f(H,P^{(t)})}{f(P^{(t)})}\left(1 - o(H)p_M - \frac{\delta(H)}{l-1}p_R P_{R,v}\right)$$

  mit durchschnittlicher Qualität $f(P^{(t)})$ (Gesamtpopulation)

  bzw. $f(H, P^{(t)})$ (Individuen in $I(H) \cap P^{(t)}$)

Folgerungen

- building blocks: Schema mit
  - überdurchschnittlicher Fitness,
  - kleiner definierender und Länge
  - geringer Ordnung

  vermehren sich rasch
- durch parallele Vermehrung solcher Eigenschaften

  sollen sich komplexere überlegene Individuen bilden

Exponentielles Wachstum

- zunächst nur eine Aussage zu einem Schritt
- wäre sie in jeder Generation wahr

  $\Rightarrow$ exponentielles Wachstum der building blocks
- Aber: Annahme gilt nicht allgemein
- durchschnittliche Güte der Population nähert sich der

  durchschnittlichen Qualität des Schemas an

Weitere Kritik

- ▷ Populationen sind sehr klein
  - ⇒ Wahrscheinlichkeitsaussagen sind kritisch
- ▷ Annahme: beobachtete Qualität eines Schemas entspricht der tatsächlichen durchschnittlichen Qualität
  - ⇒ gilt nicht bei großer Gütevarianz innerhalb eines Schemas
  - ⇒ gilt nicht bei partieller Konvergenz in der Population

Entwurfsprinzipien

- ▷ wann gelten allgemein die Aussagen des Schema-Theorems?
- ▷ können Regeln für den Entwurf neuer Algorithmen abgeleitet werden?
- ▷ Entwurfsprinzipien stellen Forderungen an das Zusammenspiel Formae–Operatoren
- ▷ stellen nur eine Möglichkeit für die Arbeitsweise der Operatoren dar

Minimale Redundanz

- ▷ möglichst: eindeutige Genotyp-Phänotyp-Abbildung
- ▷ falls Redundanz unvermeidbar ist ($\exists_{A,B\in\mathcal{G}(A\neq B)}\ dec(A) = dec(B)$): redundante Punkte in selben Formae auffangen, d.h. $[A]_\zeta = [B]_\zeta$
- ▷ sonst: unterschiedlich Formae mit ähnlicher phänotypischer Wirkung hemmen die Suche

Ähnlichkeit in Formae

- ▷ Formae sollen Individuen mit ähnlicher Güte (oder phänotypischer Ausprägung) vereinigen
- ▷ insbesondere: für Merkmale mit geringer Genauigkeit
- ▷ relevante Merkmale sollen auch in mittelgroßen Populationen statistisch als „gut“ bewertet werden

Abschluss gegen den Schnitt von Formae

- Rekombination soll sinnvolle Eigenschaften unterschiedlicher Individuen kombinieren
- durchsuchte Hyperebenen sollen sich während der Suche verfeinern
- Schnitt miteinander verträglicher Formae soll wieder eine Forma bilden

Verträglichkeit der Formae

- Entwurfsprinzip für Rekombination
- gemeinsame Eigenschaften der Eltern sollen immer auf Kinder übergehen
- $\forall_v \forall_{A,B \in v} \forall_{\xi \in \Xi}\ R^\xi(A,B) \in v$
- bisher durch die Suche Erreichtes wird dadurch erhalten
- insbesondere auch bei der Rekombination eines Individuums mit sich selbst

Übertragung von Genen

- eine Rekombination übertragt Gene oder phänotypische Allele (Formae V mit minimaler Genauigkeit), wenn für jedes neue Individuum, die Eigenschaften auf ein Elternteil zurückgeführt werden können
- $\forall_{A,B} \forall_{\xi \in \Xi} \forall_{v \in V}\ R^\xi(A,B) \in v \ \Rightarrow\ (A \in v\ \vee\ B \in v)$
- spricht für einen kombinierenden Operator
- andernfalls: Rekombination nimmt implizite Mutationen vor

Verschmelzungseigenschaft

- kompatible Formae in zwei Individuen sollen sich immer durch die Rekombination miteinander verbinden lassen
- $\forall_{v,v'(v \bowtie v')} \forall_{A \in v} \forall_{B \in v'} \exists_{\xi \in \Xi}\ R^\xi(A,B) \in v \cap v'$
- effektive Suche erfordert die Möglichkeit, alle kombinierbaren Eigenschaften zu kombinieren

Gesamtkonzept

- ▷ Entwurfsprinzipien weisen der Rekombination eine strikte Rolle des Zusammenführens von existierenden Eigenschaften zu
- ▷ Erreichbarkeit von allen Punkten im Suchraum muss durch Mutation gewährleistet werden
- ▷ sonst: zu starke Abhängigkeit von der Anfangspopulation

Eine täuschende Funktion

Genotyp	Güte
000	3.0
001	2.0
010	2.0
100	2.0

Genotyp	Güte
110	1.0
101	1.0
011	1.0
111	3.5

- ▷ durchschnittliche Güte der Schemata?

Weitere Kritik am Schema-Theorem

- ▷ *EA besitzen die Fähigkeit sich auf vielversprechende Regionen im Suchraum zu konzentrieren*
- ▷ im Schema-Theorem: Regionen $\equiv$ Schemata
  vielversprechend $\equiv$ mit überdurchschnittlicher Güte
- ▷ setzt implizit voraus:
  überdurchschnittliche Güte vererbt sich tatsächlich weiter
- ▷ fehlende Voraussetzung: Korrelation der Gütewerte zwischen Eltern und Kind

Price-Theorem: Rahmen

- ▷ wir betrachten EA (GA) als dynamisches System
  $p'_A = \sum_{B,C\in\mathcal{G}} T(A \leftarrow B, C) \frac{f(B)f(C)}{\bar{f}^2} p_B p_C$
- ▷ $p_X =$ Häufigkeit eines Individuums X
- ▷ $T(A \leftarrow B, C) =$ Wahrscheinlichkeit A aus B und C zu erzeugen

- ▷ Betrachten im weiteren eine qualitative Bewertung von einzelnen Individuen $Q : \mathcal{G} \to \mathbb{R}$ sowie von Populationen $\bar{Q} = \sum_{A\in\mathcal{G}} Q(A) p_A$

Price-Theorem

- Es gilt für zwei Eltern B und C:

  $\bar{Q} = \bar{E} + Cov\left[E[T(B,C)], \frac{f(B)f(C)}{\bar{f}^2}\right]$ mit

- erwartete Bewertung der Nachkommen:

  $E[T(B,C)] = \sum_{A\in\mathcal{G}} Q(A)T(A \leftarrow B,C)$

- durchschnittliche erwartete Bewertung in der Population:

  $\bar{E} = \sum_{B,C\in\mathcal{G}} E[T(B,C)]p_B p_C$

- Kovarianz zwischen Elterngüte und Kindbewertung:

  $$Cov\left[E[T(B,C)], \frac{f(B)f(C)}{\bar{f}^2}\right] = \sum_{B,C\in\mathcal{G}} E[T(B,C)]\frac{f(B)f(C)}{\bar{f}^2}p_B p_C - \bar{E}$$

Bewertungsfunktionen

- Häufigkeit eines Schemas H:

  $$Q_H(A) = \begin{cases} 1, & A \in I(H) \\ 0, & \text{sonst} \end{cases}$$

  $\Rightarrow$ führt zu einer Variante des Schema-Theorems

- Betrachtung eines Güteschwellwerts w:

  $$Q_w(A) = \begin{cases} 1, & A \in f(A) \succ w \\ 0, & \text{sonst} \end{cases}$$

- führt zu folgendem Theorem

Rolle der Rekombination

- Angenommen es gilt die Übertragung der Gene: vom Vater $g \in G = P^{\{1,\ldots,l\}}$, von der Mutter $\hat{g} = \{1,\ldots,l\} \setminus g$:

  $$R_g(B,C) = \begin{cases} B_i, & i \in g \\ C_i, & \text{sonst} \end{cases}$$

- sei p_g die Wahrscheinlichkeit, dass $R\ g$ erzeugt
- Dann gilt: $T(A \leftarrow B,C) = \sum_{g\in G} p_g R_g(B,C)$
- g bestimmt Menge von Schemata:

  $$H(g) = \{H_1 \ldots H_l \mid H_i \in \{0,1\} \text{ falls } i \in g,\ H_i = * \text{ sonst}\}$$

„Fehlendes“ Schema-Theorem

- $$\bar{Q}'_w - \bar{Q}_w = \sum_{g\in G} p_g\, Cov\left[Q_w(R_g(B,C)), \frac{\bar{f}_{\hat{g}}(B)\bar{f}_g(C)}{\bar{f}^2}\right] - \sum_{g\in G} p_g \sum_{B\in H(\hat{g}),C\in H(g)} (p_{R_g(B,C)} - p_B^{(\hat{g})} p_C^{(g)}) (Q_w(R_g(B,C)) - \bar{Q}_w)\frac{\bar{f}_{\hat{g}}(B)\bar{f}_g(C)}{\bar{f}^2}$$
- mit $p_B^{(\hat{g})} = \sum_{C\in H(g)} p_{R_g(B,C)}$ und $p_C^{(g)} = \sum_{B\in H(\hat{g})} p_{R_g(B,C)}$
- und Grenzgütewerten der Schemata

  $$\bar{f}_{\hat{g}}(B) = \frac{1}{p_B^{(\hat{g})}} \sum_{C\in H(g)} f(R_g(B,C))p_{R_g(B,C)}$$

  $$\bar{f}_g(C) = \frac{1}{p_C^{(g)}} \sum_{B\in H(\hat{g})} f(R_g(B,C))p_{R_g(B,C)}$$

Folgerungen

- eine positive Kovarianz zwischen den Gütewerten der Schemata und den besseren erzeugten Kindindividuen beeinflusst die Güteentwicklung positiv
- nicht alle Schemata werden verarbeitet, sondern nur diejenigen mit $p_g > 0$
- Schemata treten immer in komplementären Paaren auf

Folgerungen

- Gilt $p_{R_g(B,C)} > p_B^{(\hat{g})} p_C^{(g)}$, dann zerstört R mehr Instanzen der Kombination als dass neue geschaffen werden

  Gilt zusätzlich, dass die Instanz überdurchschnittliche Güte besitzt $Q_w(R_g(B,C)) > \bar{Q}_w$

  Dann hat die Rekombination einen negativen Einfluss auf die Güteentwicklung

  Positiver Einfluss wenn $p_{R_g(B,C)}$ unterrepräsentiert ist

Folgerungen

- die Rekombination kann dadurch optimiert werden, dass p_g so angepasst wird, dass diejenigen g mit sehr positivem Wert

  $$Cov\left[Q_w(R_g(B,C)), \frac{\bar{f}_{\hat{g}}(B)\bar{f}_g(C)}{\bar{f}^2}\right] - \sum_{B\in H(\hat{g}),C\in H(g)} \left(p_{R_g(B,C)} - p_B^{(\hat{g})} p_C^{(g)}\right) \left(Q_w(R_g(B,C)) - \bar{Q}_w\right) \frac{\bar{f}_{\hat{g}}(B)\bar{f}_g(C)}{\bar{f}^2}$$

  eine höhere Wahrscheinlichkeit erhalten