

Evolution im Computer - “künstliche Ameisen” im PC

von

Martin Reiche

1. Einleitung	3
2. Die Evolutionsaufgabe	3
2.1. Die Welt.....	4
2.2. Die Ameise	4
3. Die Evolutionsverfahren.....	6
3.1. Wichtige Konzepte.....	6
3.2. Lösung per Zufall?	6
3.3. Grundsätzlicher Ablauf des Evolutionsalgorithmus	7
4. Evolutionsexperimente	10
5. Deutung der Ergebnisse	12
6. Mathematische Betrachtung des Evolutionsprozesses	13
7. Zum Thema Zufall	17
8. Artificial Life - Leben im Computer?	18
9. Mögliche Forschungsrichtungen	18
10. Literatur	20

1. Einleitung

Im Buch "Artificial Life" von Stephen Levy {1} werden Evolutionsexperimente beschrieben, welche - auf Supercomputern ausgeführt - zu recht interessanten Ergebnissen führen. Dies hat meine Neugier geweckt, einmal selbst zu untersuchen, ob mit vertretbarem Aufwand Evolution auch im Home-Computer "funktioniert".

Wie man sehen wird, ist dies der Fall, und man gewinnt grundlegende Einsichten in die Dynamik überschaubarer Evolutionsprozesse.¹

2. Die Evolutionsaufgabe

Bei der genetischen Programmierung versucht man, problembezogene Datensätze in Richtung auf ein vorgegebenes Ziel per Evolution, sprich durch Reproduktion, Selektion und Mutation schrittweise zu optimieren. Dabei fällt dem Zufall eine wichtige Rolle zu.

Das gewählte Beispiel hat mit Absicht einen möglichst anschaulichen Charakter, obwohl dies natürlich nicht wesentlich für die genetische Programmierung ist. Es ähnelt den "künstlichen Ameisen" auf dem John Muir Trail {1, S.165ff} {2}. "Echte" Ameisen markieren einen Weg mit Duftstoffen (Pheromonen), welche ihre Artgenossen als Orientierungshilfe nutzen, dem unsichtbaren Pfad zu folgen.

Die Fähigkeit, aus einer mehr oder weniger losen Verteilung von Marken eine Spur zu lesen und ihr zu folgen, sei als Evolutionsaufgabe definiert.

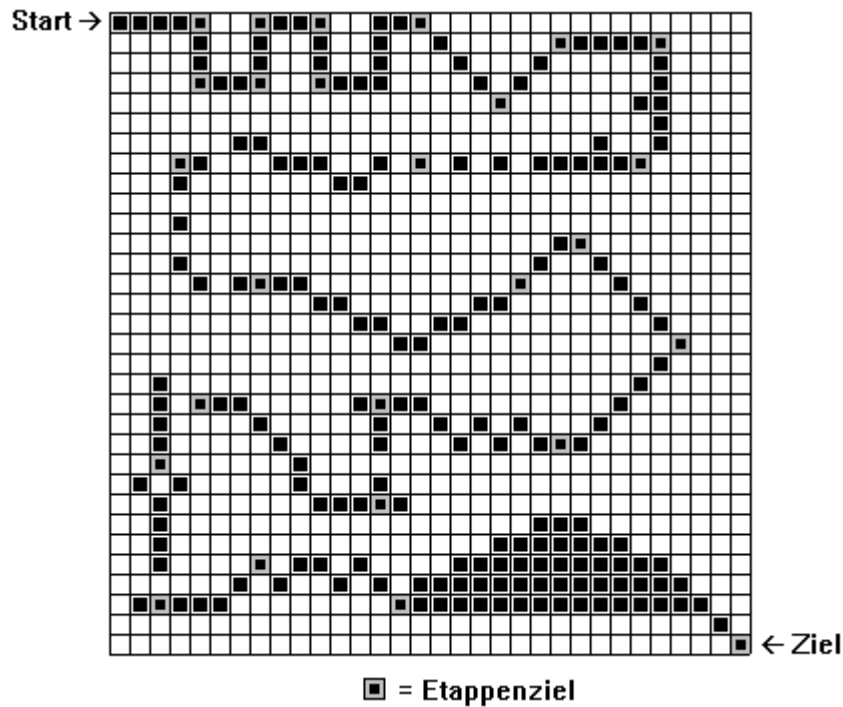
Um das Projekt einfach zu halten, dient im Experiment eine Folge von markierten Feldern auf einem karierten Feld als Spur. Die zunächst völlig ahnungslosen "Ameisen" sollen über viele Generationen hinweg "lernen", diesem Pfad zu folgen. Die zu optimierende Datenstruktur ist also sozusagen das Programm der "Ameisen", welches sie erfolgreich zum Ziel führt. Zu Beginn der Evolution ist der Programmspeicher mit rein zufälligen Regeln gefüllt. Nach und nach soll sich dann ein geeignetes Programm "von selbst" herausbilden.

Wie man feststellt, werden in diesem Papier der Anschaulichkeit halber Begriffe aus dem Bereich des Lebendigen herangezogen. Dies ist zwar bei unseren "Ameisen" lediglich metaphorisch gemeint - im Computer existiert ja kein Leben im eigentlichen Sinne - doch in anderen Fällen wird eine strukturelle Verwandtschaft zu Lebensprozessen geradezu angestrebt. Dies wird im weiteren Verlauf sichtbar. Man achte also auf die Wörter und ihre Bedeutung im gegebenen Zusammenhang! Auf Anführungsstriche wird darum ab jetzt der Lesbarkeit halber verzichtet.

¹Vom selben Autor existiert ein Microsoft Windows Programm namens "EvoLab", welches erlaubt, auch ohne Programmierkenntnisse bzw. -aufwand solche Prozesse zu studieren.

2.1. Die Welt

Besagte Ameisen bewegen sich auf einem zweidimensionalen Feld von Zellen, welche entweder markiert oder unmarkiert sein können. Die markierten Zellen bilden einen Pfad, welchem die Ameisen folgen sollen, siehe folgendes Bild:



Das Feld, sprich die Ameisenwelt, ist als Torus ausgeführt: Verläßt eine Ameise das Feld am rechten (linken, oberen, unteren) Rand, so erscheint sie augenblicklich am linken (respektive am rechten, unteren, oberen).

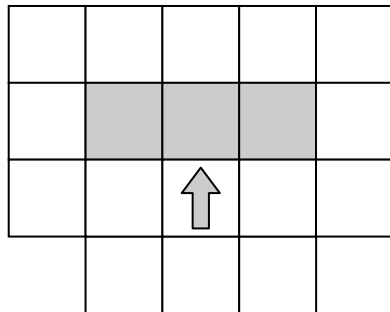
2.2. Die Ameise

Wie ist nun solch eine Ameise realisiert? Sie besteht aus

- einem Sinnesorgan, mit dem sie vorausliegende Felder wahrnimmt,
- einem inneren Zustand - eine Art Gedächtnis - , welcher die Werte 0, 1, 2 oder 3 annehmen kann,
- einem Programmspeicher mit Platz für 32 Regeln, sowie
- einem Automaten, welcher aus dem aktuellen Zustand und der Sicht anhand der Regeln den nächsten Zug ableitet.

Die Aufgabe der Evolution wird es sein, den Programmspeicher mit passenden Regeln zu belegen, sodaß diese die Ameise erfolgreich bis zum Ziel führen

Wie gesagt ist die Ameise beschränkt sehfähig: Sie kann die drei nächsten Felder in Bewegungsrichtung einsehen und feststellen, ob diese markiert sind. Im folgenden Bild markiert der Pfeil die momentane Position und Bewegungsrichtung. Die schraffierten Felder zeigen das Sehfeld.



Wie man leicht nachzählt, kann die Ameise $2^3 = 8$ verschiedene Bilder erkennen. Ihre schrittweise Fortbewegung arbeitet der Automat dann nach besagten Regeln ab, die in einer Tabelle oder Matrix festgelegt sind:

		Sicht							
Zustand	0	↑ 1	↑ 1	↑ 0	↑ 1	↑ 3	↑ 2	↑ 3	↑ 0
	1	→ 0	↑ 1	↑ 0	↓ 3	↑ 0	→ 2	← 3	→ 2
	2	← 2	↓ 0	↑ 0	↑ 1	→ 3	↓ 3	↑ 0	← 3
	3	← 0	↑ 0	← 1	↑ 3	← 1	↑ 1	↑ 0	↓ 0

Die Spalte, welche dem aktuellen Sehmuster entspricht und die Zeile des aktuellen Zustandes bestimmen eine Zelle in der Matrix und damit eine Regel. Diese hat zwei Komponenten: Die neue (relative) Bewegungsrichtung (Pfeil) und den neuen Zustand. Die Ameise zieht nun genau ein Feld weiter in die angegebene Richtung, konsultiert wieder die Matrix, und so weiter. In der Matrix bedeuten ein Pfeil nach oben "geradeaus", nach unten "zurück", nach links "links" und nach rechts "rechts".¹

Der innere Zustand hat zunächst keinerlei "Bedeutung". Er wird aber später der Ameise dienen, sich zu orientieren. "Bedeutung" ist hier also auch Ergebnis der Evolution!

¹Im Programm wird neben dem Zustand auch die Richtung in Zahlenform codiert: 0 = geradeaus, 1 = links, 2 = zurück, 3 = rechts. Die Regeln in obiger Matrix sind Ergebnis eines Evolutionsprozesses, wobei eine Schattierung bedeutet, daß diese Regel beim Passieren des Trainingsparcours nicht benutzt wird.

Betrachten wir die ersten Schritte auf dem Pfad: Ausgangssituation für die Testläufe ist das erste Feld, die Anfangsrichtung ist rechts, der Zustand mit 0 initialisiert¹. Für die ersten drei Schritte reicht eine einzige Regel: Wenn Du das Muster □■□ siehst, bleibe weiter im Zustand Null und gehe geradeaus. Auch wenn Du □■■ siehst, kannst Du noch weiter geradeausgehen und das erste Etappenziel² erreichen. Danach schaust Du aber ins Leere: Das Bild ist □□□ und Du solltest jetzt wissen, daß Du nun rechts abbiegen mußt. Zum Beispiel hättest Du bei □■■ in den Zustand 1 gehen können. Dort könnte Dir eine passende Regel sagen: Wenn Du □□□ siehst, und Du bist in Zustand 1, biege rechts ab. Somit bekommt der Zustand 1 die Bedeutung: "Abzweig nach rechts gesehen", er dient der Ameise sozusagen als Gedächtnis vorangegangener Sichten. Es leuchtet ein, daß dieser Abbiegevorgang ohne Erinnerung an vorher Gesehenes kaum zu bewältigen ist.

Mit Überlegungen dieser Art ließe sich nun die Ameise auch von Hand programmieren. (Mit EvoLab möglich!)

3. Die Evolutionsverfahren

3.1. Wichtige Konzepte

Ziel der Evolution ist die Herausbildung eines geeigneten Programmes, d.h. der Inhalt des 32 Regeln umfassenden Programmspeichers, welcher den Lauf einer Ameise bestimmt. Die Gesamtheit der 32 Regeln nennen wir Genom. (Die Ameise und ihr Genom werden im folgenden oft synonym benutzt). Mehrere Ameisen oder Individuen, die sich miteinander entwickeln, bezeichnen wir als Population. Während der Evolution sterben Individuen, doch es werden auch fortwährend neue geschaffen: Die Individuen einer Population, welche zu einer Zeit geschaffen werden, nennen wir eine Generation. Werden einzelne Regeln zufällig verändert, sprechen wir von Mutation.

Den Grad der Eignung bezeichnen wir als die Fitness eines Genoms. Sie wird hier durch einen Testlauf ermittelt: Der Testkandidat wird auf die Startposition des Testparcours gesetzt. Insgesamt 26 Positionen des Parcours stellen sogenannte Etappenziele dar, die die Ameise nacheinander innerhalb einer jeweils begrenzten Anzahl von Schritten erreichen muß (Siehe Bild oben). Kommt eine Ameise ans Ziel, besitzt sie die maximale Fitness von 26, erreicht sie nicht einmal die erste, ist ihre Fitness gleich Null. Kommt sie z.B. nach Passieren der Etappe 13 nicht rechtzeitig bei 14 an, wird die Simulation abgebrochen, ihre Fitness wird mit 13 festgehalten. Somit ist die Fitness definiert als die Nummer der zuletzt erfolgreich absolvierten Etappe bzw. des letzten erreichten Etappenziels (Start = Etappe 0). Mathematisch läßt sich die Fitnessfunktion f definieren, welche jedem denkbaren Genom für einen gegebenen Parcours genau einen Fitnesswert zuordnet.

3.2. Lösung per Zufall?

Man könnte geneigt sein, die Programmierung ganz dem Zufall zu überlassen, vielleicht landet man ja einen Treffer. Doch halt, wieviel "Programme" gibt es eigentlich? Da genau

¹Wegen der Gleichberechtigung aller Orte, Richtungen und Zustände bedeutet dies aber keine Einschränkung der Allgemeinheit

²Die Etappenziele sind für die Ameise nicht sichtbar, sie dienen nur der Bewertung der Fitness.

16 mögliche Regeln existieren um jede der 32 Zellen zu belegen sind es 16^{32} , das ist eine Zahl mit 38 Nullen! Viel zuviel, um jemals von allen Rechnern der Welt in Millionen von Jahren ausprobiert werden zu können. Doch vielleicht gibt es ja auch ungeheuer viele mögliche Lösungen? Läßt man den blinden Zufall ans Werk, ergibt sich bei 40 Millionen zufällig gewählten Genomen folgendes Bild:

Erreichte Fitness	Anzahl
0	39.079.477
1	890.745
2	27.785
3	1.056
4	151
5	0
6	0
7	624
8	162
9 und größer	0
Summe	40.000.000

Für das Auftreten der Fitness 1 ließe sich mit etwas Fleiß eine Wahrscheinlichkeitsbetrachtung machen. Direkt angeben kann man eine untere Schranke für die erwartete Zahl: Lautet die Regel in Zeile 1, Spalte 3: "Bleibe in Zustand 0 und gehe geradeaus", so wird im Mittel jedes sechzehnte zufällig gewählte Individuum die ersten 3 Schritte geradeaus schaffen. Für den vierten Schritt paßt jede vierte Regel am Platz (1,4), nämlich eine, welche die Richtung beibehält. Somit sollten bei 40 Mio. Versuchen mehr als $40.000.000 / (4 * 16) = 625.000$ Individuen aufgrund dieser Regeln über eine Fitness von 1 verfügen. Daß noch andere Regeln zur ersten Etappe führen, zeigen die restlichen 265.745 Individuen.

3.3. Grundsätzlicher Ablauf des Evolutionsalgorithmus

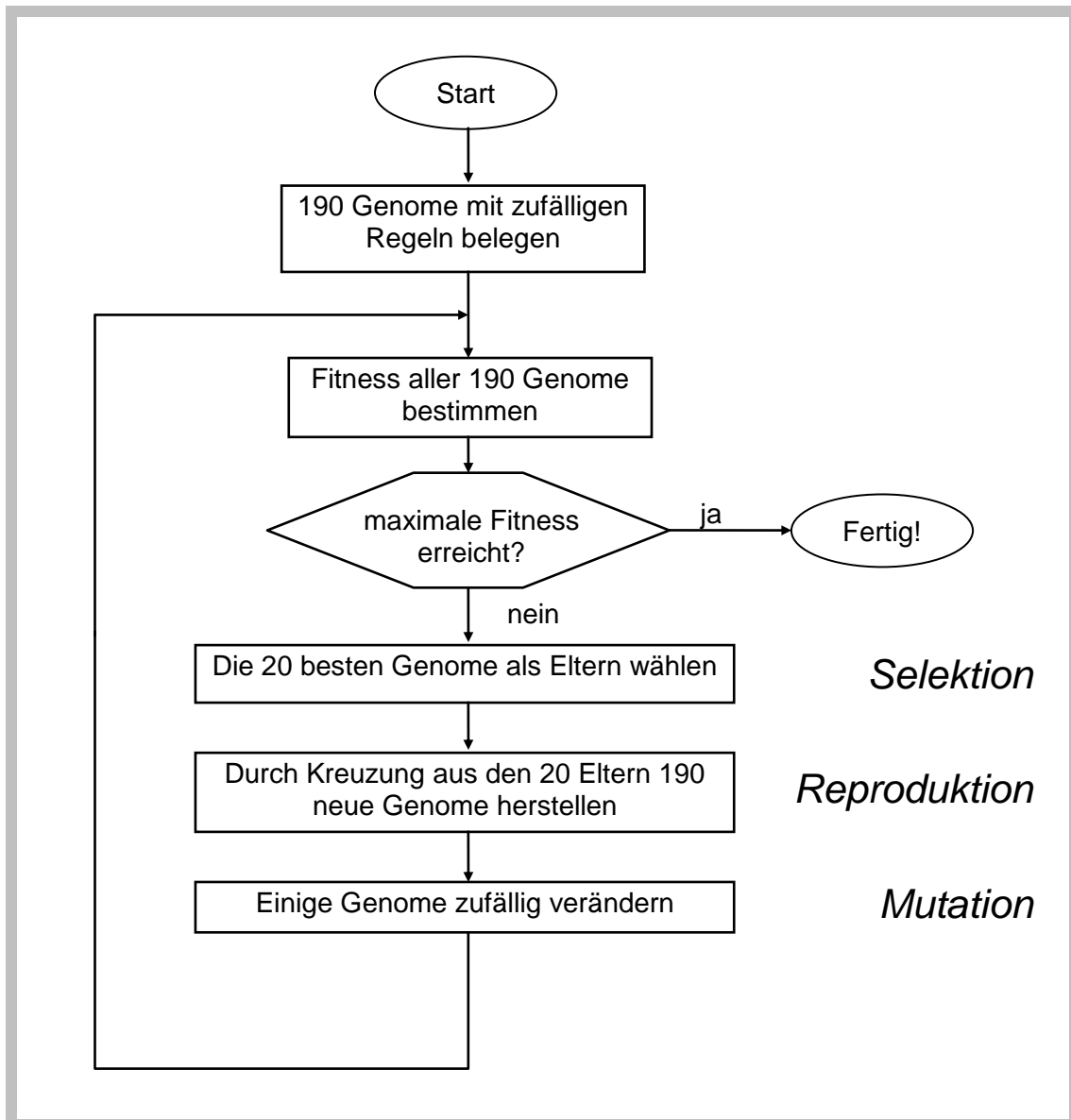
Das Evolutionsprogramm funktioniert nach folgendem Schema: Man erzeugt eine größere Population von z.B. 190 Ameisen per Zufall und prüft ihre Fitness. Die meisten dieser Ameisen werden es nicht einmal bis zu ersten Etappe schaffen, es besteht ja für sie zunächst keinerlei Zusammenhang zwischen markierten Zellen und einer Spur! Doch sind in der Regel auch einige darunter, die so weit kommen¹. Aus der Gesamtpopulation werden nun die besten 20 ausgewählt und miteinander gekreuzt. Wenn aus jeder Paarung eine neue Ameise hervorgeht, ist die neue Generation mit $n*(n-1)/2$ sprich $20*19/2=190$ Individuen wieder vollständig. Auf diese Population wirkt nun eine zufällige Mutation indem z.B. 500 der $190*32$ d.h. etwa 8% der Regeln zufällig verändert werden. Diese Population wird dann wieder dem Fitness-Test unterworfen, und so weiter.

¹Siehe obige Wahrscheinlichkeitsbetrachtung!

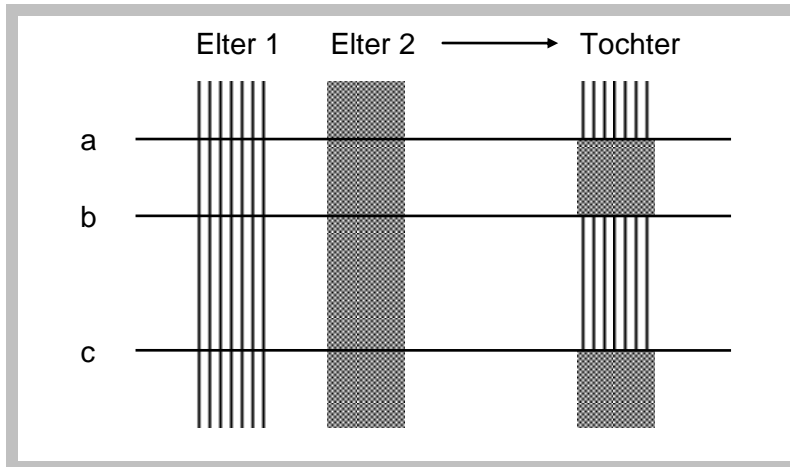
Immer dann, wenn in der Population ein Genom mit einer höheren Fitness als bis dato beobachtet auftritt, sprechen wir von einem Evolutionsschritt.

Zum "Auswürfeln" einer Regel dient ein im Computer bereitgestellter "Zufalls"zahlen-generator. Dieser liefert eine Serie von Zahlen, die ihr Muster erst nach vier Milliarden Werten wiederholt. Startet man den Generator mit einem sogenannten Saatwert (engl. seed), so lassen sich verschiedene Serien abrufen.

Grafisch kann man das Verfahren als Flußdiagramm so darstellen:



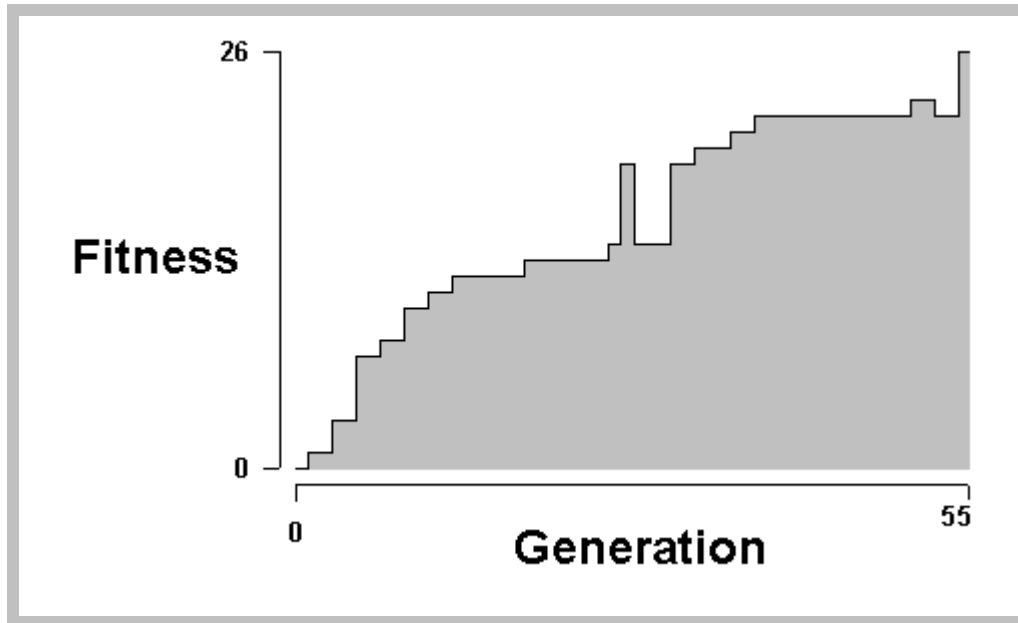
Die Kreuzung von Genomen geschieht nach Art des auch in der Natur vorkommenden "Crossing-over". Dazu wird zunächst jede Elternmatrix in je eine einzelne Spalte überführt, indem man ihre Regeln - gelesen von links nach rechts und oben nach unten - untereinander schreibt. Danach wählt man eine Anzahl Überkreuzungspunkte zufällig und erzeugt ein Tochtergenom, indem man abwechselnd Regeln vom ersten und zweiten Elternteil übernimmt. Dies zeigt das folgende Bild:



Abschließend kann man die Matrix durch den umgekehrten Prozeß wiederherstellen.

4. Evolutionsexperimente

Im Laufe der Generationen zeigt sich in der Tat eine wachsende Fitness. Das Bild zeigt eine recht stürmische Entwicklung: in nur 55 Generationen wird die maximale Fitness von 26 erreicht.



Natürlich kann aufgrund der "blinden" Mutation und des Crossing-Over eine einmal erreichte Fitness auch wieder verlorengehen - die Fitnesskurve sinkt dann, wie oben ersichtlich, zwischenzeitlich wieder ab.

Weniger rasante Evolutionsläufe - das sind die meisten! - unterscheiden sich darin, daß z.T. über viele Generationen hinweg keine Verbesserung zu verzeichnen ist. Aus diesem Grunde wird im Programm ein Evolutionslauf auch nach z.B. 200 Generationen abgebrochen, falls sich die maximale Fitness nicht einstellt.

Einen wichtigen Evolutionsparameter stellt die Anzahl der selektierten Eltern dar: Neben den oben erwähnten 20 wurden in einem Versuch auch einmal nur das beste Genom ausgewählt d.h. die Elternzahl zu 1 gesetzt. Natürlich findet dann keine Neukombination der Gene statt. Das Crossing-over entfällt, d.h. bevor die Mutation ausgeführt wird, sind alle 190 Genome identisch

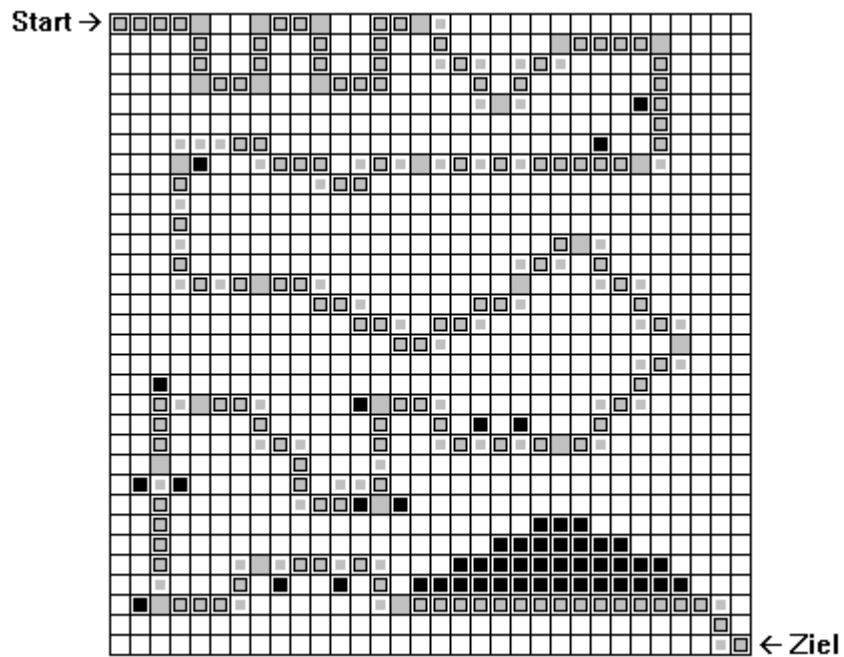
Bei 1000 Evolutionsläufen ergab sich folgendes Bild:

	Anzahl der Eltern	
	1	20
Erfolgreiche Evolutionsläufe	66	151

Die "Gen-Rekombination" bringt also in diesem Versuch einen merklichen Vorteil.

Grundsätzlich bemerkenswert bei diesen Versuchen ist, daß der Evolutionsalgorithmus offensichtlich in keinerlei Beziehung steht zu der "Bedeutung" und inneren Struktur des Genoms - und trotzdem funktioniert!

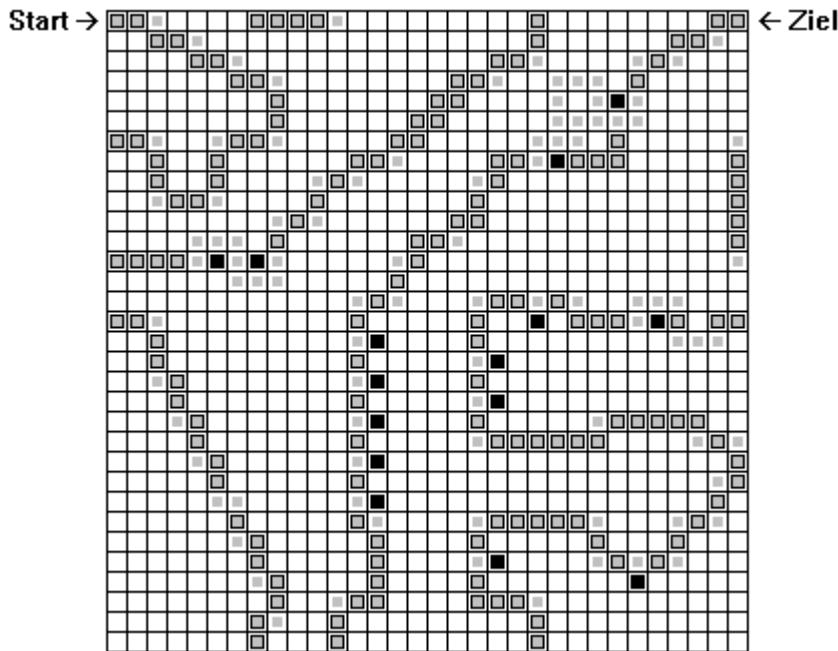
Das folgende Bild zeigt, wie sich eine evoluierte Ameise auf dem Parcours bewegt:



Die von der Ameise betretenen Felder sind mit einem kleinen grauen Quadrat markiert.

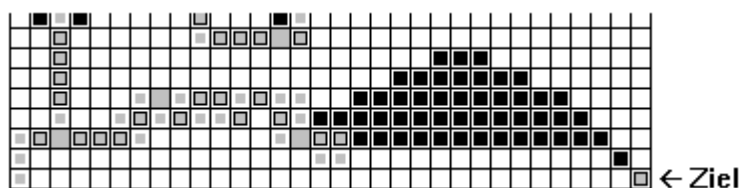
5. Deutung der Ergebnisse

Was bedeuten diese Evolutionserfolge? Können wir nun sagen, die Ameisen hätten gelernt, einer Spur zu folgen? Sicher ist eigentlich nur, daß ihr Programm sie mit der vorgegebenen Verteilung von Marken rechtzeitig zu allen Etappenzielen führt. Was passiert zum Beispiel, wenn man sie mit einem ganz anderen Pfad konfrontiert? Einen Testlauf zeigt folgendes Bild:



Man sieht, auch dieser wird erfolgreich absolviert - wenn auch mit einer "Ehrenrunde" oben rechts. Dies ist aber beileibe nicht selbstverständlich: Den meisten evoluierten Genomen gelingt es nicht, dem alternativen Pfad bis zum Ende zu folgen! Immerhin ist damit aber gezeigt, daß durch die Evolution mehr erreicht wird, als das Auswendiglernen eines bestimmten Pfades. Würden denn 32 Regeln überhaupt ausreichen, um den Trainingspfad "abzubilden"?

Auch Ameisen mit maximaler Fitness sind nicht perfekt. Was heißt aber perfekt überhaupt? Sie sind so gut, wie wir sie gezüchtet haben, und keinen Deut besser. Dies mag ein Beispiel verdeutlichen. Folgendes Bild zeigt den Lauf einer Ameise mit der maximalen Fitness von 26 am Ende des Parcours:



Obwohl sie unseren formalen Anforderungen genügt, folgt sie in der letzten Etappe nicht der Spur, sondern erreicht das Ziel, indem sie umkehrt und über den linken Rand auf das Etappenziel stößt.

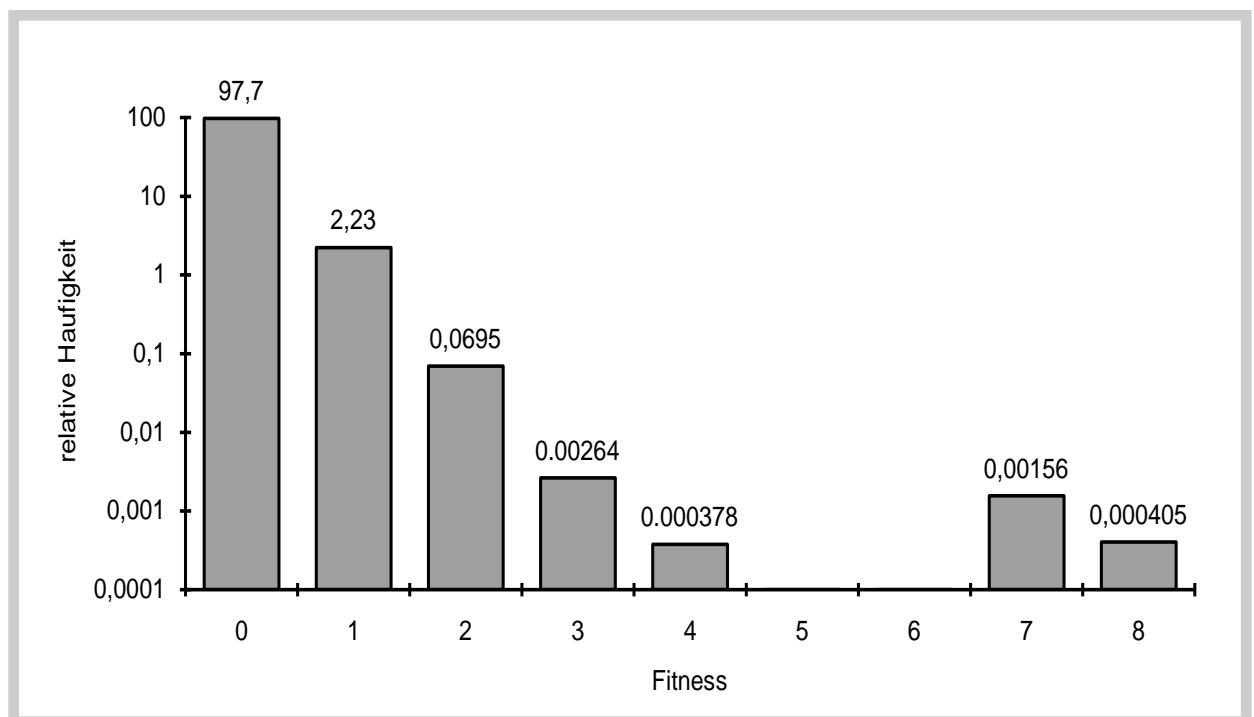
6. Mathematische Betrachtung des Evolutionsprozesses

Die Evolutionsaufgabe lässt sich mathematisch wie folgt betrachten: Die 32 Speicher für Regeln spannen einen 32-dimensionalen Hyperraum (Sequenzraum $\{3\}$) auf, in dem die Fitnessfunktion f jedem Satz von Regeln $r_1 - r_{32}$ genau einen Wert zuordnet:

$$0 \leq f(r_1, r_2, \dots, r_{32}) \leq 26$$

Für jeden Parcours existiert genau eine Funktion f . Bildlich gesprochen erzeugt f eine Wertelandschaft $\{3\}$ mit Bergen und Tälern, wobei allerdings die Höhe nicht die dritte Dimension ausmacht, sondern die 33ste. Übertragen auf dieses Bild bedeutet Maximierung der Fitness das Besteigen des höchsten Berges.

Was können wir nun über diese Wertelandschaft sagen? Der oben erwähnte Versuch, per Zufall zu einer Lösung zu gelangen, liefert eine Stichprobe über die Höhenverteilungen der Landschaft, die als Schätzung so dargestellt werden kann:



97.7% aller denkbaren Genome haben die Fitness 0 - der weitaus größte Teil der "Landschaft" ist also durch eine monotone "Tiefebene" gekennzeichnet. Kann man umgekehrt die Anzahl aller Genome maximaler Fitness abschätzen? Einen Anhaltspunkt für eine untere Schranke liefern die Evolutionsergebnisse: Aus einem speziellen

Evolutionsexperiment resultierten 151 verschiedene Genome¹. Wenn also 151 verschiedene Lösungen gefunden wurden, wie groß ist dann wahrscheinlich die Gesamtzahl der Lösungen mindestens? Dieses Problem ist folgendem äquivalent: Gegeben sei ein Urne mit einer unbekanntem Anzahl N Kugeln. (N entspricht der unbekanntem Zahl der Lösungen) Nun werden nacheinander Kugeln gezogen, markiert und wieder zurückgelegt. Wenn man dieses n mal durchführt, ohne eine markierte Kugel ein zweites Mal zu ziehen, wie groß ist dann N mit 95% Wahrscheinlichkeit mindestens? Die Wahrscheinlichkeit p , n -mal nacheinander eine nichtmarkierte Kugel zu ziehen, ist

$$p(n, N) = \prod_{k=2}^n \left(1 - \frac{k-1}{N}\right) \quad \text{mit } 1 < n \leq N+1$$

Durch Iteration läßt sich für $p = 0,05$ und $n = 151$ die untere Schranke zu $N = 3831$ berechnen.

Für die obere Schranke ist die Rechnung einfacher: Sei N wieder die Anzahl der Optimalgenome, $G = 16^{32}$ der Gesamttraum aller Genome und n die Anzahl der Stichproben aus diesem. Dann ist die Wahrscheinlichkeit p , bei n Versuchen kein Optimalgenom zu finden

$$p = \left(1 - \frac{N}{G}\right)^n$$

Weil N , wie oben gezeigt, sehr viel kleiner als G ist, kann man annähern

$$p \approx \left(1 - \frac{nN}{G}\right)$$

Hat man bei n Versuchen nichts gefunden, so muß mit 95% Wahrscheinlichkeit ($p = 0.05$) gelten:

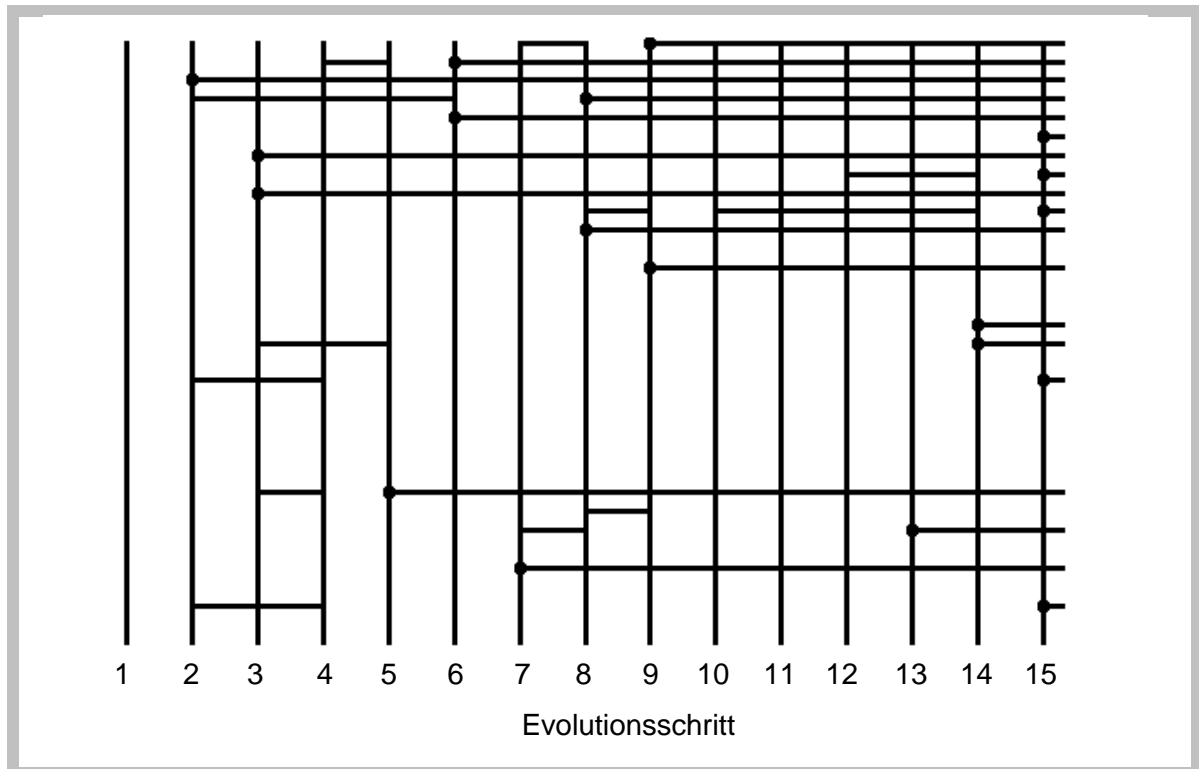
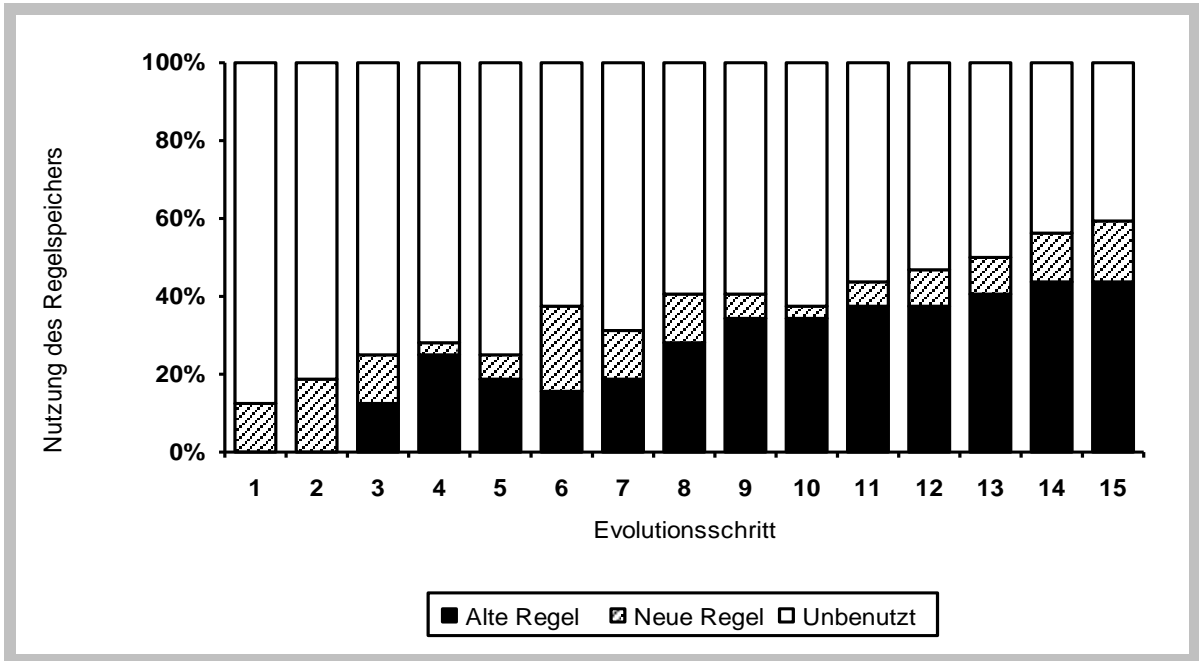
$$N < (1 - p) \frac{G}{n}$$

- was immer noch einen Wert von $8,1 * 10^{30}$ ergibt.

Gemessen an der Vielzahl der möglichen $3,4 * 10^{38}$ Regelkombinationen hat es die Evolution aber immerhin wiederholt geschafft, die sprichwörtliche Stecknadel im Heuhaufen zu finden. Dies ist nur durch bestimmte lokale Eigenschaften der Funktion f zu erklären: Hat man eine bestimmte Fitness von z.B. 16 erreicht, so ist die Wahrscheinlichkeit groß, in der Nachbarschaft, d.h. durch kleinere Modifikationen der Regeln, ein Genom mit einer Fitness von 17 oder mehr zu finden. Der Evolutionsalgorithmus ist dann in der Lage, diesen Ort aufzuspüren.

¹Allerdings ließe sich einwenden, daß trotz der Verschiedenheit der Genome die "Logik" vieler gleich ist, weil sich aus einer gegebenen Lösung durch Permutation der 4 Zustände ja weitere gleichwertige ableiten lassen. Darum ist scheint es angeraten, die Wege auf dem Trainingspfad miteinander zu vergleichen. Nach ersten Ergebnissen scheint dies aber keinen signifikanten Unterschied zu machen: Von 59 untersuchten Genomen war kein Pfad wie der andere.

Das folgende Bild zeigt diesen schrittweisen Neuerwerb von Regeln während eines Evolutionslaufes: Jeder Balken charakterisiert die 32 Regeln in jedem Evolutionsschritt (Def. siehe 3.3): Unbenutzte Regeln sind weiß, aus dem letzten Schritt übernommene Regeln schwarz und neu erworbene Regeln gestreift dargestellt.

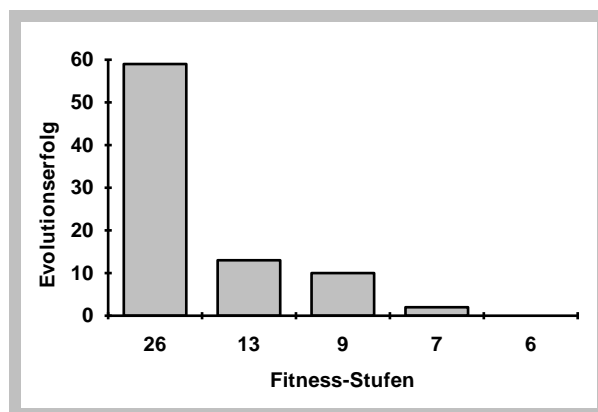


Die zweite Darstellung zeigt ebenfalls die Historie desselben Genoms: Die senkrechten Linien bezeichnen die Vorfahren aus jedem Evolutionsschritt. Das erste Genom befindet sich links, das mit der maximalen Fitness rechts außen. Die Orte der Regeln auf dem senkrechten Balken reihen sich von oben nach unten. Zwischen zwei Genomen bezeichnet eine waagerechte Linie die Identität zweier (jeweils auch benutzter) Regeln.

Eine durchgezogene Linie vom rechten Rand endet bei dem Genom weiter links, bei dem die zugehörige Regel in der Evolutionsgeschichte zum ersten Male auftrat und dann nicht mehr verlorenging. Dieser Endpunkt ist gesondert markiert. Man sieht, das Optimalgenom hat Regeln aus vielen Schritten ererbt. Auch kann es vorkommen, daß einmal erworbene Regeln später durch andere ersetzt werden. Die waagerechte Linie endet dann nicht am rechten Rand.

Manfred Eigen beschreibt das so {3}: Bei der "evolutiven Gipfeltour ... gibt es für den Evolutionsprozeß im hochdimensionalen Sequenzraum mit großer Wahrscheinlichkeit eine kontinuierliche Sprungroute zum Gipfel der Wertelandschaft," was dadurch begründet ist, daß man an jedem Punkt eine große Anzahl von möglichen Bewegungsrichtungen findet und "nirgendwo weiter als n Schritte vom höchsten Gipfel entfernt ist (wobei n die Sequenzlänge ist)". Die Sequenzlänge entspricht in unserem Falle der Anzahl der Regeln.

Experimentell läßt sich diese Deutung bestätigen. Man reduziert dazu z.B. die Stufenzahl der Fitnessfunktion. Genome mit benachbarten Fitnesswerten liegen nun nicht mehr so nahe beieinander, die besagte lokale Eigenschaft von f wird reduziert. Programmtechnisch ist dies sehr einfach zu erreichen: Nach Bestimmung der Fitness eines Genoms wird nicht die zuletzt erreichte Etappe (genannt "eta"), sondern für eta kleiner 26 nur der ganzzahlige Anteil von eta/n an den Evolutionsalgorithmus übergeben. Mit geringerer Zahl der Fitness-Stufen zeigt sich dann ein stark abnehmender Evolutionserfolg bei 200 Evolutionsläufen, wie man folgender Grafik entnehmen kann:



Der Evolution fällt es zunehmend schwer, zum nächsten höherwertigen Punkt im Sequenzraum zu gelangen, denn nun müssen mehrere Regeln verändert werden, um eine höhere Fitness zu erreichen. Bezogen auf den Parcours bedeutet das: Das Vorankommen und Erreichen der nächsten Etappen wird erst deutlich später belohnt - der Evolution wird sozusagen der "Kompaß" entzogen.

Daß die lokale Eigenschaft auch für $n=1$ nicht an allen Punkten gegeben ist zeigt die Tatsache, daß manche Evolutionsprozesse über Hunderte von Generationen keinen Fortschritt erzielen. Alle Individuen sitzen dann sozusagen auf einem isolierten, kleineren Hügel der Wertelandschaft, dessen "Fangbereich" sie nicht verlassen können: Der Weg zu

einem höheren Berg bedeutet einen Marsch durch die Niederung. Jeder Ausreißversuch in dieser Richtung wird aber sofort Opfer der Selektion.

Diese Beobachtungen geben einen Hinweis darauf, unter welchen Bedingungen genetische Algorithmen erfolgversprechend anwendbar sind:

1. Eine Lösung des Problems ist gegeben durch einen Satz von Parametern.
2. Die meisten (besser: alle) Parameterkombinationen sind gültige Alternativen.
3. Die Überprüfung eines Parametersatzes auf Tauglichkeit ist nicht zu rechenintensiv.
4. Man muß eine Fitnessfunktion aufstellen können, welche Parametersätzen, die näher an der Lösung liegen, auch höhere Werte zuweist.

Evolutionsalgorithmen scheinen also auch da anwendbar zu sein, wo bekannte Optimierungsverfahren versagen: In unserem Problem sind die Regeln entlang einer Dimension nicht geordnet - Zwar ist jede Regel als Zahlenwert codiert, doch ist f nicht "stetig": Erhöht man den Code einer Regel um eins, so ergibt sich meist eine ganz andere Struktur des "Programms" und die Fitness des Genoms kann sich drastisch ändern. Evolution kann also nicht nur Fließkommazahlen fein justieren wie z.B. herkömmliche Gradientenverfahren sondern auch neue Strukturen schaffen!

Aus diesem Grunde der Nicht-Stetigkeit hinkt auch der oben angeführte Vergleich einer Werte"landschaft".

Überflüssig zu erwähnen, daß sich die Ergebnisse des vorliegenden punktuellen Experiments natürlich nicht einfach auf andere Bereiche übertragen lassen!

7. Zum Thema Zufall

Die Ausgangsgenome sowie die spontanen Mutationen werden im Computer mit einem sogenannten "Zufallszahlengenerator" produziert, der nach einem strengen Rechenverfahren arbeitet. Das äußert sich darin, daß ein gegebenes Programm vollkommen deterministisch abläuft - Es kommt immer dasselbe heraus, egal wie oft man das Programm startet. Wie kann man da noch von Zufall sprechen? Zufällig heißt so viel wie unvorhersehbar, unerwartet. Für unser Programm reicht es also augenscheinlich, daß keine Verwandtschaft oder Ähnlichkeit besteht zwischen der Funktion des Zufallszahlengenerators und dem Rest des Programms. Für den Skeptiker bliebe nachzuprüfen, ob der Erfolg des Evolutionsprogrammes mit "echten" Zufallszahlen aus dem Würfel, einer Lotto-Maschine oder einem Geigerzähler ausbleibt. Ich bin sicher: Er wird es nicht!

Vielleicht stellt der Leser in seinem Denken einen bedeutenden Rollenwechsel des Zufalls fest: Vom Gegenspieler, der verhindert, daß spontan Ordnung entsteht - zum Helfen, welcher die Ordnung erst ermöglicht und dabei bestimmt, welche Ordnung zur Ausprägung kommt.

Natürlich ist denkbar, daß ein deterministischer Suchalgorithmus schneller zum Ziel führt als die ungerichtete Mutation. Das schmälert aber nicht die Bedeutung der Tatsache, daß der Evolution auch eine unkorrelierte, "zufällige" Veränderung zum Erfolg genügt.

8. Artificial Life - Leben im Computer?

Anhand eines anschaulichen Beispiels aus dem Bereich "genetische Programmierung" haben wir gezeigt, wie Begriffe aus der Biologie - der Lehre vom Leben - durchaus sinnvoll auf "tote" Computerprogramme und ihre Ergebnisse anwendbar sind. Dies mag zunächst etwas seltsam klingen, denn Elektronenrechner und Lebenwesen sind ja grundsätzlich verschiedener Natur. Es kommt aber, wie man sieht, ja auch gar nicht auf das ausführende Organ der zu betrachtenden Programme an, sondern nur auf die Ausführung an sich. Im Prinzip sind sie auch mit Würfel, Papier und Bleistift (und Kopf!) durchführbar - wenn auch etwas zeitaufwendiger.

Für manchen mögen die Ergebnisse überraschend erscheinen. Konnte man vorher wissen, ob der beschriebene "Versuchsaufbau" wirklich "Evolution" produzieren oder nur blind vor sich hindümpeln und nichts "sinnvolles" zuwege bringen würde? Weder unsere Alltagserfahrung noch Vorstellungskraft gibt ja konkrete Hinweise auf die Dynamik solcher Systeme - und daher muß man es ausprobieren. Was man daraus lernen kann: Evolution funktioniert - auch in nichtbelebten Systemen wie dem vorliegenden Programm.

Man neigt vielleicht dazu, folgende Unterscheidung zu machen: Folgt die Ameise "wirklich" einer Spur, oder erscheint dies nur so? Hat die Ameise denn überhaupt eine Vorstellung davon, was eine Spur ausmacht? Ich glaube, diese Fragen enthüllen nur ein Problem unserer Sprache und damit unseres Denkens. Welche "Vorstellung" hat denn eine Biene von einer Blüte, die Katze von der Maus oder der Mensch von der "Natur"? Die aktuelle Antwort der Wissenschaft ist simpel: Kommt der Träger einer Vorstellung damit in der Welt zurecht, dann erfüllt die Vorstellung ihren Zweck. Stirbt dagegen ein Individuum aufgrund mangelhafter Vorstellung von der Welt oder scheitert eine Theorie im Experiment, dann war diese eben mit der Welt nicht vereinbar. Ein Mehr an Wirklichkeit zu fordern macht keinen Sinn. Ergo folgt die Ameise wirklich der Spur.

9. Mögliche Forschungsrichtungen

- Man könnte die ausgelegte Spur etwas kompakter und schwieriger gestalten und die maximalen Etappenzeiten verkürzen: Würden die so gezüchteten Ameisen sich auf anderen Strecken dann besser bewähren? Oder müßte man dazu den Parcours deutlich vergrößern, und mit allen Finessen bestücken etc. Wie lang würde dann die Evolution bis zur maximalen Fitness dauern? Würden immer mehr Regeln benötigt? Stagniert die Evolution, sobald ein Großteil des Regelspeichers "verbraucht" ist?
- Auch die Fitnessfunktion ließe sich differenzieren: So könnte die benötigte Schrittzahl mit eingerechnet werden und ebenso die Anzahl der verwendeten Regeln. Geht die Konzentration einher mit einer Robustheit und verbesserten Leistung auf fremden Strecken? Was hat die Reduktion auf wenige Regeln mit "Abstraktion" zu tun?
- Finden sich solche Ameisen besser zurecht, welche den nächsten Schritt auch abhängig machen von der zuletzt gewählten Richtung, oder müßte man ihre "Sehleistung" verbessern? Sie könnten auch Markierungen auf den schon besuchten Plätzen hinterlegen, um nicht an einer unübersichtlichen Stelle in der Spur zurückzulaufen (wie beobachtet)?
- Gibt es technische Anwendungen z.B. in der digitalen Bildverarbeitung? Etwas höher gezüchtete "Ameisen" könnten Spuren folgen, Unterbrechungen finden, Konturen erfassen, Umfänge und Flächen ermitteln etc.

- Wie kann man die Anzahl der Genome maximaler Fitness genauer ermitteln?

10. Literatur

- {1}: Steven Levy: "Artificial Life", New York 1992.
Paperback unter ISBN 0-679-74389-8.
Deutsche Ausgabe: "KL - Künstliches Leben aus dem Computer", Droemer Knaur 1993, ISBN 3-426-26477-3

- {2}: Christopher G. Langton et. al. (editors): "Artificial Life II" (Proceedings of the Workshop on Artificial Life Held February, 1990 in Santa Fe, New Mexico) Addison Wesley 1992, ISBN 0-201-52570-4. - ISBN 0-201-52571-2 (paperback)

- {3}: Manfred Eigen: "Stufen zum Leben" Piper Taschenbuch

- {4}: David E. Goldberg: "Genetic Algorithms in Search, Optimization, and Machine Learning", Addison Wesley 1989

- {5}: Eberhard Schöneburg et.al. "Genetische Algorithmen und Evolutionsstrategien", Addison Wesley 1994

- {6}: Bernd-Olaf Küppers: "Der Ursprung biologischer Information", Piper 1990

- {7}: Paul Kußmaul: "Der König unter den Algorithmen - Der Evolution in die Karten geschaut". Unveröffentlichter Bericht, 1995

- {8}: Martin Reiche: "Evolution im Computer - Ein erstes Projekt". Unveröffentlichter Bericht, 1994