Small summery

Boris Prochnau

9. Juni 2014

Inhaltsverzeichnis

1 First meeting (14.04): Model of interest

These are some main points of the model we are using for the simulation. They are the product of the first meeting with Loren Coquille and Martina Baar.

- We have N traints and $(X_{i=1,\dots,N})$ are the amount of members of the traits.
- We use a constant mutation rate that does not depend on the traits μ.
- but the other rates are depending on traits and change with time (by as the amount of traitmembers rises)
  - b_i, [intrinsic birth-rate]
  - d_i, [intrinsic death-rate]
  - $\left(\sum_{j=1}^{N} \frac{c_{\cdot,j}}{K} x_j\right)_i$, [competition death]
  - $\mu\left(\frac{b_{i-1}+b_{i+1}}{2}\right)$, [extrinsic birth-rate]
- We use 3 groups of PPP's where every group represents an event.
  - $N_t^{b_i}$
  - $N_t^{d_i}$
  - $N_t^{\bar{c}_i}$

  where $\bar{c} = \left(\sum_{j=1}^{N} \frac{c_{\cdot,j}}{K} x_j\right)_i$. Therefore we get 3N processes running that compete about the first one occurring with each one triggering an death/birth for an trait.
- With respect to the fact that the distribution of the N_t changes with the size of the population we can think of an reset of the parameters of N_t due to the fact that the increments are exponentially distributed and therefore are memoryless (markov property)
- The use of coloring one trait was (only $\rightarrow$ ask later) to explain the superposition of the PPP associated with the simulation to extract the occured event from the PPP's

  $$\rightarrow N_t^{total_i} = N_t^{b_i+d_i+\bar{c}_i} = N_t^{b_i} + N_t^{d_i} + N_t^{\bar{c}_i}$$

  Therefore we would decide:

  $$X_i^{total} = \begin{cases} \text{coloring d} & \text{with prob. p} \\ \text{coloring b} & \text{with prob. q} \\ \text{coloring } \bar{c} & \text{with prob. 1-p-q} \end{cases} \tag{1}$$

  with:

  $$p = \frac{d_i}{d_i + b_i + \bar{c}_i} \qquad q = \frac{b_i}{d_i + b_i + \bar{c}_i} \qquad 1 - p - q = \frac{\bar{c}_i}{d_i + b_i + \bar{c}_i}$$

2 Second meeting (22.04): Pseudocode

First thing to mention is that the Pseudocode snippets are marked as „Algorithm“, but in fact they are just functions and I don't know yet how to change the name in this particular LaTeX environment. Also I use a „=“ instead of „←“ because I think it provides better readability.
The following Pseudocode will contain many function-calls. Every function can be identified as a verb, and objects or local variables are nouns. There will also be one Boolean „isBorn“ which is an adjective.
Generally the functions have no return values because in this situation we tell functions to do something (not ask). Means we pass a task to them. There will be no requests for return values.
One other thing to mention is that the function names can get long, but they will (or should) always say what the function does. Not more or less.
The function-body should always follow a rule called „pretty much what you expected“. Some functions do violent this rule slightly. They are usually more than 6 lines and will be changed later.
First we start with the main Step of the Algorithm and continue with further explaining of the used functions. When a new function appear, it will be explained directly below it. Also we seperate this section in 3 parts, each one is dedicated to one of the functions used in this following EvolutionStep():

Algorithm 1 EvolutionStep()

Require: -
Ensure: A full evolution Step happened
1: calculateEventRates();
2: sampleEventTime();
3: changeATrait();

This function does a full evolution step. First thing to do is to calculate the Event rates with „calculateEventRates“, than we sample an exponential time with our current rate parameters with „sampleEventTime“ and finish the Step with changing a Trait with „changeATrait“.
Most functions-calls should not need to be explained, that's why I leave this explanations out in following comments.
This function will be improved later with not calculating the Rates in every new step, rather than updating them with the previous changes.

2.1 Calculating total-event rates

We will encounter 2 classes called „Trait“ and „Events“. They are so called Data Classes and usually mainly store data.
They are visible at all time and can be used like global variables.
In this first part of calculating total-event rates we only need to introduce the

„Trait“ class which will be used as an array of traits „Trait[$i \leq n$]“ with n beeing the maximum number of traits. An Trait Object has attributes that can be accessed through a dot operator like „Trait[i].Members“ what would be „the Members of Trait i“.
This is a listing of all attributes in one Trait Object:
class Trait

- BirthRate
- TotalBirthRate
- DeathRate
- TotalDeathRate
- TotalTraitRate
- TotalEventRate - [static]
- CompDeathRate[i][j] - [static]

Here is much work with Superposition of the PPP included. We summarize the sum of birth-rates from members within a trait to the „TotalBirthRate“ (mutation included), same for „TotalDeathRate“ (competition included).
The „TotalTraitRate“ is the summed “TotalBirthRate“ and „TotalDeathRate“ and means the total rate of events for a specific trait.
„TotalEventRate“ and „CompDeathRate“ are static, this means that they are the same for all initialized Trait objects. They (all items from the array) access the same variable for this. „TotalEventRate“ is the same as usual and „CompDeathRate“ is the competition-matrix with according rates.

Algorithm 2 calculateEventRates()

Require: -
Ensure: All (total)Rates will be set

```
1: for i=0 to n-1 do
2:     calculateTotalDeathRateOf(i)
3: end for
4: calculateTotalBirthRates(0);
5: calculateTotalEventRate();
```

In the next function we will manipulate attributes of the Trait Objects. The mutation will be accessed without and Traitindex: „Trait.Mutation“ because it is static and therefore the same for all Objects.

Algorithm 3 calculateTotalBirthRates(StartIndex: i)

Require: int i
Ensure: Total birthrate of Trait „i“ will be set (recursively)
1: Trait[i].TotalBirthRate = (Trait[i].Members)·(Trait[i].BirthRate)
2: **if** $i < n - 1$ **then**
3: calculateTotalBirthRates(i+1)
4: Trait[i].TotalBirthRate += $\frac{Trait.Mutation}{2} \cdot Trait[i+1].TotalBirthRate$
5: **end if**
6: **if** $i > 0$ **then**
7: Trait[i].TotalBirthRate += $\frac{Trait.Mutation}{2} \cdot Trait[i-1].TotalBirthRate$
8: **end if**

In Algorithm 3 in line 3 is used recursion, because this improves the calculation speed a lot, although it slightly makes code less intuitive.

Algorithm 4 calculateTotalDeathRateOf(TraitIndex: i)

[H]
Require: int i
Ensure: Total deathrate of Trait „i“ will be set
1: Trait[i].TotalDeathRate = 0;
2: addTotalIntrinsicDeathRateOf(i);
3: addTotCompetitionDeathRateOf(i);

Algorithm 5 addTotalIntrinsicDeathRateOf(TraitIndex: i)

1: Trait[i].TotalDeathRate = (Trait[i].DeathRate) · (Trait[i].Members)

Algorithm 6 addTotalCompetitionDeathRateOf(TraitIndex: i)

1: **for** j=0 **to** n-1 **do**
2: Trait[i].TotalDeathRate += (Trait.CompDeathRate[i,j])·(Trait[j].Members);
3: **end for**

Algorithm 7 calculateTotalEventRate()

Require: -
Ensure: Current Totaleventrate is set
1: **for** i=0 **to** n-1 **do**
2: Trait[i].TotalTraitRate = Trait[i].TotalBirthRate
+ Trait[i].TotalDeathRate;
3: Trait.TotalEventRate += Trait[i].TotalTraitRate;
4: **end for**

2.2 Sampling the next event-time

Here will appear a, not yet mentioned, object that will not be explained further, called Dice. The Dice Object will provide a uniform or exponential random Variable.

Algorithm 8 sampleEventTime()

Require: -
Ensure: First ringing Eventclock has been sampled
1: double Parameter = Trait.TotalEventRate;
2: double newEvent = this.Dice.RollExpDice(Parameter);
3: Events.EventTimes.push(newEvent);

Here we use Dice.RollExpDice(λ) to get $X \sim exp(\lambda)$. The same is possible for Dice.RollUnifDice(λ) to get $X \sim Unif[0, \lambda]$.

2.3 Changing a trait

Here we will work with the actual Events taking place. For this purpose i introduce the Events class like before the Trait class:
class Events

- Dice
- EventTimes[i]
- ChosenTrait[i]
- isBirth[]

The EventTime, ChosenTrait and isBirth are so called vectors. They are dynamic containers and we need to push an item into them to append it to the last entry.

Algorithm 9 changeATrait()

Require: -
Ensure: make a change to the Population with current Parameters
1: choseTraitToChange();
2: choseEventType();
3: executeEventTypeOnTrait();

Algorithm 10 choseTraitToChange()

Require: -
Ensure: Trait is chosen for changing

```
1: double Parameter = Trait.TotalEventRate;
2: double HittenTrait = Dice.rollUnif(Parameter);
3: for i = 1 to n-1 do
4:    if HittenTrait ≤ Trait[i].TotalEvent then
5:       this.ChosenTrait.push(i);
6:       break;
7:    end if
8:    HittenTrait -= Trait[i].TotalEvent;
9: end for
```

Algorithm 11 choseEventType()

Require: -
Ensure: Decision for Birth or Death is made

```
1: int i = Events.ChosenTrait.lastentry();
2: double EventType = Dice.rollUnif(Trait[i].TotalTraitRate);
3: if EventType ≤ Trait[i].TotalBirthRate then
4:    Events.isBirth.push(true);
5: else
6:    Events.isBirth.push(false);
7: end if
```

Algorithm 12 executeEventTypeOnTrait()

Require: -
Ensure: Chosen event will occur on chosen trait

```
1: if isBirth then
2:    Trait[ChosenTrait.lastentry()] += 1;
3: else
4:    if ChosenTrait.Members > 0 then
5:       Trait[ChosenTrait.lastentry()] -= 1;
6:    end if
7: end if
```

List of Algorithms

3 Thrid and Fourth meeting (Di 20.05 and 13.05): Graphical plans - afterwards

Zunächst habe nicht meine bisherigen Fortschritte präsentiert. Dazu gehörte ... Außerdem haben wir die graphischen Ansprüche an das Programm besprochen. Dabei wurden folgende Ergebnisse erzielt:

- Man sollte eine gute Möglichkeit haben die Parameter im Programm einsehen zu können und auch eine gute Möglichkeit Daten zu laden.
- Was besonders wertvoll ist, wäre eine Möglichkeit neue Instanzen zu generieren. Dabei hat sich jedoch das Problem gestellt dass es keine einfache Konsole mit simplem einseitigen Output und Input gibt, da man während einer Instanzgenerierung auch weiterhin mit dem Hauptfenster interagieren könnte. Das Problem wurde etwas umständlich über eine kontrollierte Deaktivierung der anderen Interfaceelemente gelöst.

Zur Darstellung des Graphen wurde folgendes gesagt:

- Es ist keine "Filmfunktion" gewünscht. Stattdessen sollte man einen Plot generieren und in diesen hereinzoomen und mit der Maus ziehen können.
- Außerdem ist eine Legende gewünscht die eine Übersicht über alle Graphen bietet.
- Um die Graphen später vergleichen zu können sollte es eine Möglichkeit geben den Graph zu speichern.

4 Fifth meeting (Fr 30.05): Improvements - afterwards

An diesem Tag habe bloß einige Änderungen vorgestellt und sichergestellt dass alle Funktionen ordnungsgemäß laufen:

4.1 Graphische Änderungen

Das waren hoffentlich alle optisch merklichen Änderungen:

- Ich habe eine Checkbox hinzugefügt die eine Option für sinnvollen Abbruch bietet. Diese Funktion prüft ob sich beide Populationen in einer /*epsilon* Umgebung um den erwarteten Zustand befindet und stoppt anschließend die Iterationen. Zu Gunsten der Übersicht werden noch mal 1/3 der Anzahl bisher gemachten Iterationen angefügt.
  Zu diesem Zeitpunkt ist $\epsilon = \frac{5}{\sqrt{K}}$.

- Es wurde ein Anzeigefehler behoben der für Geburts- und Todesraten immer nur die Populationsgrößen der Eigenschaften angezeigt hat.

- Auf dem Plotfenster habe ich ein Label mit den gemachten Iterationen hinzugefügt. Weiterhin findet sich dort jetzt auch ein Feld wo man einen Dateinamen eingeben kann der einem gespeichterten Bild dienen soll welches man mit dem neuen "save Image" Button erstellen kann. Das Bild wird zur Zeit als .png und .pdf gespeichert.

- Die x-Achsenbeschriftung lautet nun nicht mehr Millisekunden sondern "Time". Grund dafür ist dass ich nicht sicher bin welche Zeit unsere exponentielle Uhr ausgibt. Jedoch habe ich sie der Übersicht halber nicht mehr durch 1000 geteilt, da bei den gewünschten Instanzen ausreichend große Zeiten angenommen werden.

- Zuletzt wurde noch bei der Legende der für die k-te Erwartete Eigenschaft "k. Expected" zusätzlich der konvergierte Wert angezeigt "k. Exp: 1.048".

4.2 Interne Änderungen

- Ich habe mehrere Arten der Datenspeicherung ausprobiert. Zunächst habe ich die Daten alle nacheinander in eine "Storage" Datei geschrieben um schließlich einzeln die Daten mit "addData" zum Graphen hinzuzufügen. Das schien mir jedoch an der "addData" Stelle und am Datastream zu aufwendig. Danach bin ich auf eine alternative mit "Jumped Steps" umgestiegen. Die Jumped Steps basieren darauf das maximal 10.000.000 Punkte gespeichert werden und auch wenn viel mehr Iterationen gemacht werden, so werden nur Äquidistante 10 mio Stützstellen geplottet. Der Gedanke dahinter war die nicht Unterscheidbarkeit des Graphen für das menschliche Auge.

- Ich habe eine "GraphClass" hinzugefügt die alle Daten für das "PlotWindow" vorbereitet, so dass das Fenster nur noch darauf zugreifen muss. Damit habe ich eine Brücke zwischen den Fenstern und dem Plot geschlagen.
- Eine der wichtigsten Änderungen ist das Hinzufügen des QThreads. Dieser gewährleistet dass der Plot die Fenster nicht einfrieren lässt. Die Berechnung kann für mehr als 1mio Punkte bereits merkliche Zeit (¿3s) beanspruchen.
- ...to do...

4.3 bekannte Bugs

Es gibt hier noch einen Bug der mit der Beschriftung der x-Achse zu tun hat. Ich weiß noch nicht wie man ihn reproduziert, aber er könnte die Achse um einen großen Faktor zu groß skalieren. Der Plot weißt jedoch das geplante Verhalten auf und dieser Bug tritt nur bei ŕeplots; also dann auf wenn ein neuer Plot den alten überschreibt.

5 Sixth meeting (03.06): Introducing TSS

Einführung in TSS

- Das erste war das Einfügen einer Tabelle oder Grafik die einen guten Überblick über die Fitness der Population bietet. Diese gibt Aufschluss darüber ob TSS ordnungsgemäßes Verhalten zeigen wird. Allgemein ist es aber auch gut zu sehen welche Population wie stark auf eine andere wirkt, daher vielleicht schon ohne TSS sinnvoll.
  Um eine bessere Übersicht von ungewünschtem Verhalten zu zeigen werden Coexistenzen rot markiert. Später wird noch eine weitere grünliche Färbung angezeigt die eine Verdrängungswahrscheinlichkeit darstellt. Dieses Feature wird in beiden Programmen vorhanden sein.
  Die verwendete Formel zur Berechnung der Fitness war

  $$f_{x,y} = b_x - d_x - c_{x,y} \cdot \bar{n_x}$$

  mit:

  $$\bar{n_x} = \frac{b_x - d_x}{c_{x,x}}$$

  Demnach wird auf der Diagonale stets eine Null erwartet. Bei dieser Rechnung tritt jedoch eine Subtraktion auf die Auslöschung verursacht. Zu diesem Zweck habe ich die dargestellte Zahl auf 14 stellen runden lassen.

- Ein besseres Verständnis von TSS wäre nützlich. Dazu wurde mir aus [2] S.1135 empfohlen. Leider habe ich es bisher noch nicht geschafft es durchzuarbeiten. Es wird auf den nächsten Wochenplan kommen.

- Außerdem wurde die Skalierung der Mutationswahrscheinlichkeit mit K angepasst so dass TSS realistisch laufen kann. Dabei wurde der Faktor $\frac{1}{K^{1.5}}$ gewählt weil es den notwendigen Schranken die in [2] S.1135 beschrieben werden genügt. Ein weiterer Punkt den man optimieren kann wäre es eine Option zur Eingabe eines eigenen Exponenten für K wäre eine mögliche Erweiterung.

- In diesem Zustand ist es bereits Möglich einen TSS plot zu erzeugen. Jedoch wird die Zeit im Equilibrium tatsächlich berechnet was viel unnötige Rechenarbeit impliziert. Später solle eine lineare Interpolation bis zur ersten Mutation gemacht werden sobald das Equilibrium erreicht wird.

5.1 Bekannte Bugs

Es ist ein Bug bisher übrig. Bisher kenne ich noch nicht den Ursprung, aber er äußert sich (meistens) durch einen Programmabsturz nachdem ich eine kleinere Instanz lade und anschließend TSS1 bzw eine größere Instanz lade. Der Plot wurde dabei noch nicht ausgeführt.

6 Seventh meeting (10.06): Bug fixing and interpolation of Equilibrium

6.1 Was ist geplant?

Geplant ist:

- Beheben eines Fehlers beim Laden der Instanzen. Dieser wurde im letzten Kapitel beschrieben.
- Eine lineare Interpolation der Population im Equilibrium. Die Funktion dazu wurde bereits geschrieben.
- Korrekte Ausarbeitung eines geeigneten Abbruchkriteriums für TSS Prozesse.
- Punkte an denen eine Mutation stattgefunden hat sollen idealerweise hervorgehoben werden. Z.B. durch kreuze oder ähnliches.
- Am Besten wäre noch eine Skalierung der Zeitachse während des Equilibriums.
- Die Fitnessmatrix sollte bisher eine Bandmatrix sein weil es nur Mutationen zum Nachbarn gibt.

6.2 Was habe ich gemacht?

Konkret:

- Bisher habe ich eine Neudefinition der "isNear()" Funktion gemacht. Diese gibt es nun als "isNearDimorph()", "isNearMonomorph()" und "isNearTSS()". Diese tun was man vom Namen her erwartet, wobei "isNearTSS()" prüft die Nähe zum Equilibrium bzw. es prüft ob die Population bereits auf das Equilibrium getroffen ist (mit Rücksicht auf Auslöschung bis zu einer Genauigkeit von 10^{-10}) und ob die Population zur Zeit die einzig aktive ist.
- Weiterhin existiert eine neue Version der "iterateGraphPoint()" Methode. Sie heißt "iterateMutationPoint()" und soll den Zeitpunkt der nächsten Mutation der Nachbarn berechnen. Anschließend soll dieser Zeitpunkt in der Zeitlinie gespeichert werden und dort eine Mutation auf der Population ausgeführt werden.
- Diese "iterateMutationPoint()" Funktion soll anschließend von einer neuen Funktion "makeTSSIterations()" nur dann ausgeführt werden, wenn die neue "isNearTSS()" eine Bestätigung liefert. All diese Änderungen wurden in der GraphClass gemacht und werden in "generateEvolution()" verwendet.

- Ein wichtiger Zusatz für die Berechnung der Geburtenrate wurde getroffen. Dabei wurde jetzt nicht mehr zu beginn $b(i) \cdot N(i) + MutationFromNeighbors$ gerechnet, sondern $b(i) \cdot N(i) \cdot (1-\mu) + MutationFromNeighbors$. Das ist bisher nicht aufgefallen weil μ für gewöhnliche dimorphe Prozesse immer Null war.

- Eine sehr wichtige Änderung sollte die Berechnung des monomorphen stabilen Zustands gelten. Wenn eine positive Mutationswahrscheinlichkeit vorliegt, so sinkt die Arteigene Geburtenrate um die Mutationswahrscheinlichkeit. Somit berechnet sich das Equilibrium im TSS durch

$$\frac{b(x) \cdot (1-\mu) - d(x)}{c_{x,x}}$$

- NearTSS wurden die Grenzen geändert, so dass es jetzt eine Umgebung von 1/K um das Equilibrium gibt statt dem genauen Treffer. Das ist notwendig, weil es sein kann dass unser Equilibrium kein Vielfaches von 1/K ist, gewährleistet aber trotzdem die gröst mögliche Nähe zum Equilibrium.

- Es hat sich ein Bug ergeben der die Zeit falsch (bzw immer mit der ersten Rate) berechnet hat. Das wurde durch ein veraltetes "static" beim Ziehen der Würfelergebnisse verursacht.

- Ein weiteres Problem wurde behoben nach dem zufolge es noch keine ordentlich Anzeige für ein TSS gab, nachdem ein Trait zum ersten mal ausgestorben ist. Die Ursache lag in der "isNearTSS()" Funktion, daher wurde diese Überarbeitet.

- Es gab ein Problem weshalb der angezeigte stabile Zustand nicht bis zum Mutationspunkt durchgezogen wurde, sondern bereits sichtbar davor stoppte. Die Ursache war dass nur die Zeit bis zur nächsten Mutation als neuer Zeitpunkt gespeichert wurde, nicht aber die iterierte Zeit.

- Ich habe einen Bug behoben der das Programm manchmal zum Absturz brachte. Dabei handelte es sich wahrscheinlich um einen Speicherfehler, die genaue Ursache kenne ich nicht, aber ich konnte den Fehler beheben indem ich die Fitness Matrix "gecleart()" habe bevor sie erneut "resize()" aufruft.

todo

- Es sollten direkte getter und setter auf die Trait und Event Members ausgeführt werden satt sie temporär zu speichern wie in "generateEvolution()".

6.3 Bekannte Bugs

- Der letzte Punkt des Graphen wurde nicht gespeichert. Außerdem sollte der Graph auch wenn nichts mehr für ihn passiert, trotzdem bis zum ende

fortgesetzt werden. Dazu sollte der letzte Zeitpunkt nocheinmal mit dem Endzustand abgespeichert werden.

- Exp in der Legende sollte gesondert geändert werden.
- Problem mit K-member K-expected und Darstellung beider.

7 Zweites Bachelorseminar

7.1 Ziel

Das Ziel meiner Bachelorarbeit ist es ein Programm zu entwickeln welches eine Simulation

7.2 Model

Das verwendete Model

7.3 Algorithmus

Die Algorithmen

7.4 Simulation

7.4.1 Aufgabenteilung und Flexibilität

Drei Module, jedes hat seine eigene Aufgabe.

7.4.2 Diagramme

Klassendiagramm
Aktivitätsdiagramm

7.4.3 Graphische Darstellung

Kein Film, zoom, Skalierung, speichern der Bilder etc...

7.5 Programmlayout

7.5.1 Laden der Parameter

Beste Option ist lesen der Parameter aus Dateien. Dabei wurde gewünscht dass kein direkter Zugriff des Nutzers auf die Dateien notwendig sein sollte um alle verwendeten Parameter anzuzeigen.

7.5.2 Erstellen einer Testinstanz

Da generell der Umgang mit Dateien und speziell auf das Programm zugeschnittenen Regeln für das Beschreiben dieser unerwünscht ist, habe ich eine Möglichkeit gesucht Dateien über das Programm mit Benutzereingabe erstellen zu lassen.

7.5.3 Präsentation des Programms

7.6 Korrektheit

7.6.1 Korrektheit des Algorithmus

7.6.2 Korrektheit der Implementation

Das viel interessantere ist die Korrektheit der Implementation. Diese ist generell mit steigender Komplexität schwerer zu prüfen (besonders bei Zufallsbedingten Simulationen). Daher habe ich das Prinzip der "Testgetriebene Entwicklung" (Test Driven Develeopment) verwendet.

7.7 Warum ist es vielleicht schwierig simple Mechanik einzupflegen?

Wie man sich vielleicht denken kann sind viele kleine Features an und für sich sehr simpel. Unglücklicherweise gibt es viele kleine solche Features und Daten die mit dem Programm verbunden wären.

8 Weitere geplante Punkte

8.1 große Verbesserungen

- Es wäre sinnvoll wenn es eine Möglichkeit gibt die angestrebten stabilen Zustände als Option auszuwählen welche angezeigt werden sollten. Dabei sollte die Beschriftung "Dimorph" oder "Monomporph" darin auftauchen.
- Es sollte das aktuell verwendete K angezeigt werden.

8.2 kleine Verbesserungen

- Items in der Tabelle sollten zentrieren
- stabiler Zustand weniger Aufdringlich (z.B. größerer Abstand zwischen Strichen und geringere Strichstärke)