One Approach for Evaluating the Ecological Sustainability of Agricultural Practices at the Landscape-scale

bmarron

16 October 2015

Avena+ Test Bed (created by Benedikt GroB)

What we'll cover

Personal motivations and bias

Some big questions and a few focused research questions

A very promising research approach

A novel methodology algorithm

Some exploratory results

Potential challenges and pitfalls

Next steps

Motivations and personal bias

No, thanks.

- currently radiative forcing is about $1.5 \mathrm{Watts} \backslash \mathrm{m}^2$
- currently we fix around 125 $\mathrm{TgN} \backslash \mathrm{yr}$
- currently extinction rates are 10-1000x background

Motivations and personal bias

Yes, please!

"Paris 2050 Project", Eco-utopian architect Vincent Callebaut

Motivations and personal bias

Maybe this but scaled up?

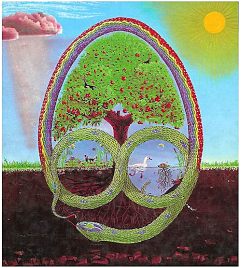
"The Garden of Abundance", Les Jardins de La Chatonniere, Ahmed Azeroual, Gardener-in-Chief

Motivations and personal bias

Rome's Eco-District, Vincent Callebaut

Food production systems that offer:
* minimal material and energetic inputs
* minimal externalized environmental and social costs
* maximal nutritional output to People and Planet
* maximal biodiversity

Motivations and personal bias



the egg of life

∗care for the Earth
∗care for the People
∗take only what you need
∗give back

Our end game

Rainbow Valley Farm, Matakana, NZ

Robust and adaptive food production systems designed at a regionally-specific, landscape scale (aka, 'sustainable agriculture')

But what does 'sustainable agriculture' actually mean?

[see handout1, please]

A few of the big questions

$*$ Does a given food production system on a landscape enhance ecological resiliency or not?

A few of the big questions

∗ Does a given food production system on a landscape enhance ecological resiliency or not?

∗ What is the preferred configuration and practice of agriculture on any given landscape?

A few of the big questions

∗ Does a given food production system on a landscape enhance ecological resiliency or not?

∗ What is the preferred configuration and practice of agriculture on any given landscape?

∗ How can the dynamics of agro-ecological systems be used to develop dichotomous or synoptic keys for the classification of agricultural practices along the continuum of ecological resliency?

A few of the big questions

∗ Does a given food production system on a landscape enhance or reduce ecological resiliency?

∗ What is the preferred configuration and practice of agriculture on any given landscape?

∗ How can the dynamics of agro-ecological systems be used to develop dichotomous or synoptic keys for the classification of agricultural practices along the continuum of ecological resiliency?

∗ What suite of pattern recognition tools can be used to logically, and ultimately causally, link ecological processes to landscape-scale patterns?

Important research questions for me

$*$ What is the relationship between landscape-scale disturbance regimes, agricultural output, and biodiversity? Which landscape mosaic is the 'most appropriate' one to maintain?

Important research questions for me

∗ What is the relationship between landscape-scale disturbance regimes, agricultural output, and biodiversity? Which landscape mosaic is the 'most appropriate' one to maintain?

∗ What is the n-dimensional vector of state variables from agro-ecological systems that is necessary and sufficient to assess the current ecological health and resiliency of a given landscape pattern and to evaluate the ecological health and resiliency of simulated landscape patterns (scenarios)?

Important research questions for me

∗ What is the relationship between landscape-scale disturbance regimes, agricultural output, and biodiversity? Which landscape mosaic is the 'most appropriate' one to maintain?

∗ What is the n-dimensional vector of state variables from agro-ecological systems that is necessary and sufficient to assess the current ecological health and resiliency of a given landscape pattern and to evaluate the ecological health and resiliency of simulated landscape patterns (scenarios)?

∗ Can a modeling environment be created that captures the spatial and temporal evolution of those critical ecological and environmental processes which are, in large measure, responsible for the patterned heterogeneity of an agro-ecological landscape?

Important research questions for me

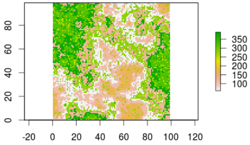
∗ What is the relationship between landscape-scale disturbance regimes, agricultural output, and biodiversity? Which landscape mosaic is the 'most appropriate' one to maintain?

∗ What is the n-dimensional vector of state variables from agro-ecological systems that is necessary and sufficient to assess the current ecological health and resiliency of a given landscape pattern and to evaluate the ecological health and resiliency of simulated landscape patterns (scenarios)?

∗ Can a modeling environment be created that captures the spatial and temporal evolution of those critical ecological and environmental processes which are, in large measure, responsible for the patterned heterogeneity of an agro-ecological landscape?

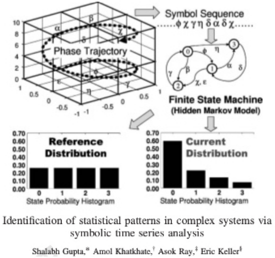
∗ What are the critical elements and processes shared by historical examples of both successful and disasterous agroecological practices that can inform our transition to sustainable agriculture?

A very promising research approach



LANDIS-II

+



Symbolic Dynamics

Why this approach?

[see handout2, please]

Why this approach?

Identification of statistical patterns in complex systems via symbolic time series analysis

Shalabh Gupta,* Amol Khatkhate,† Asok Ray,‡ Eric Keller§
The Pennsylvania State University, University Park, PA 16802, USA

(Received 27 August 2005; accepted 26 February 2006)

ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Genetics & Development

Logical and symbolic analysis of robust biological dynamics

Leon Glass[1] and Hava T Siegelmann[2]

VOLUME 90, NUMBER 10 PHYSICAL REVIEW LETTERS week ending 14 MARCH 2003

Detecting Subthreshold Events in Noisy Data by Symbolic Dynamics

Peter beim Graben[1,2,*] and Jürgen Kurths[2]
[1]*Institute of Linguistics, Universität Potsdam, P.O. Box 601553, 14415 Potsdam, Germany*
[2]*Institute of Physics, Nonlinear Dynamics Group, Universität Potsdam, P.O. Box 601553, 14415 Potsdam, Germany*
(Received 11 January 2002; published 14 March 2003)

Why this approach?

International Journal of Bifurcation and Chaos, Vol. 21, No. 12 (2011) 3465–3475
© World Scientific Publishing Company
DOI: 10.1142/S0218127411030660

SYMBOLIC DYNAMICS, COARSE GRAINING AND THE MONITORING OF COMPLEX SYSTEMS

VASILEIOS BASIOS
Interdisciplinary Center for Nonlinear Phenomena & Complex Systems, Université Libre de Bruxelles, Brussels, Belgium
vbasios@ulb.ac.be

DÓNAL MAC KERNAN
School of Physics, University College Dublin, Dublin 4, Ireland

PRL **102,** 088701 (2009) PHYSICAL REVIEW LETTERS week ending 27 FEBRUARY 2009

Symbolic Dynamics of Biological Feedback Networks

Simone Pigolotti, Sandeep Krishna, and Mogens H. Jensen
*Niels Bohr Institute and Niels Bohr International Academy, Blegdamsvej 17, DK-2100 Copenhagen, Denmark**

(Received 2 June 2008; published 26 February 2009)

PROCEEDINGS OF THE 6TH ESGCO 2010, APRIL 12-14, 2010, BERLIN, GERMANY 1

Classifying Cardiac Biosignals using Order Pattern Statistics and Symbolic Dynamics

Ulrich Parlitz, Sebastian Berg, Stefan Luther, Alexander Schirdewan, Jürgen Kurths and Niels Wessel

Why this approach?

Here's what Claire says,
"We are creating spatially explicit models based on resource availability, resource needs, and foraging scales, to develop alternative scenarios for managing the agro-natural landscape for pollination function."

– Claire Kremen, Department of Ecology and Evolutionary Biology, Princeton

[Ecology Letters, (2005) 8: 468-479]

The proposed algorithm

Level I: Build an m-dimensional vector of state variables.

Level II: Define a set of landscape reference scenarios.

Level III: Translate the landscape reference scenarios to LANDIS-II.

Level IV: Build a landscape reference scenario codebook.

The proposed algorithm

Level V: Build m-order Markov chains (bigram models) for the reference landscape scenarios.

Level VI: Define agro-ecological sustainability for a given landscape.

Level VII: Evaluate the sustainability vocabulary of extant and novel landscape designs.

The proposed algorithm

Level I: Build an m-dimensional vector of state variables.

$$\overrightarrow{KMVV} = (kmv_1, kmv_2, kmv_3, kmv_4, kmv_5, ...kmv_m)$$

Translate observable measures to an m-dimensional feature representation vector.

Level I: Build an m-dimensional vector of state variables.

Step 1 ==> Select the set of key monitoring variables (KMVs) for agro-ecosystems at the landscape scale

Many choices!

[see handout1, please]

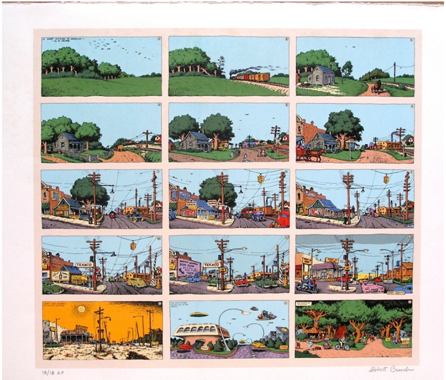
Level I: Build an m-dimensional vector of state variables.

Step 2 ==> Fix the n-dimensional vector of required empirical variables (the REVV)

==> Automate the calculation of the m-dimensional vector of key monitoring variables (the KMVV)

The proposed algorithm

Level II: Define a set of landscape reference scenarios.



R. Crumb

Level II: Define a set of landscape reference scenarios.

Step 3a ==> Define five reference systems of human impact and presence on the landscape (so-called HIP classes)

==> HIP classes A through E are defined by % agriculture, % forest, and % built environment on the landscape

==> HIP subclasses are defined by unique definitions of agriculture, forest, and built environment

Level II: Define a set of landscape reference scenarios.

HIP Class	% Ag.	% Forest	% Built Env.
A	5.0	90.0	5.0
B	20.0	75.0	5.0
C	40.0	50.0	10.0
D	80.0	5.0	15.0
E	5.0	5.0	90.0

Five reference systems of human impact and presence (HIP) on the landscape.

Level II: Define a set of landscape reference scenarios.

Step 3b ==> Define 'agriculture', 'forest', and 'built environment' for HIP subclass1

Agriculture	==	intensely managed land using standard, *agribusiness practices*
Forest	==	lightly-managed land covered in locally-appropriate vegetation
Built environment	==	land completely unavailable for photosynthetic activity (zero net primary productivity)

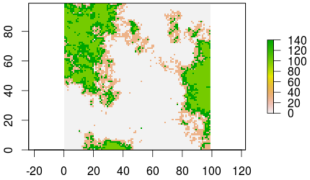
Level II: Define a set of landscape reference scenarios.

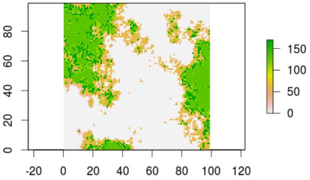
Step 3c ==> Define 'agriculture', 'forest', and 'built environment' for HIP subclass2

Agriculture	==	intensely managed land using *agro-ecological practices*
Forest	==	lightly-managed land covered in locally-appropriate vegetation
Built environment	==	land completely unavailable for photosynthetic activity (zero net primary productivity)

The proposed algorithm

Level III: Translate the landscape reference scenarios to LANDIS-II.





$\xrightarrow{Time}$

Level III: Translate the landscape reference scenarios to LANDIS-II.

Step 4a $==>$ Define the set of landscape patches for each of the reference systems defined by an HIP subclass

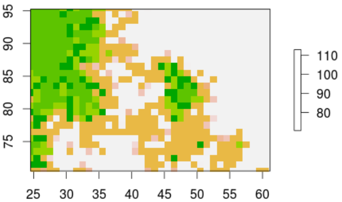
$$HIP\ subclasses = \{A1, B1, C1, D1, E1, A2, B2, C2, D2, E2\}$$

Level III: Translate the landscape reference scenarios to LANDIS-II.

Step 4b ==> Provide each patch-type with a unique moniker, a unique MapCode, a spp. list, and spp. age distribution for building LANDIS-II 'initial communities' files

```
LandisData   "Initial Communities"
>>Old jackpine oak
MapCode  0
   acerrubr 30
   pinubank 80 90
   pinuresi 110 140
   querelli 40 120 240
>> young jackpine oak
MapCode  1
   pinubank 30 50
   querelli 10 40 70
```

Level III: Translate the landscape reference scenarios to LANDIS-II.



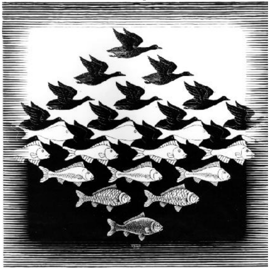

```
 initialCommunities[1,]
 [1] 5 5 6 5 5 5 4 4 3 4 4 4 5 4 4 3 3 4 5 5 5 6 6 6 6 6 6 6 6 6 5 4 3 4 4 4 5
[38] 5 5 4 4 4 2 2 1 1 1 1 1 2 2 1 1 1 2 2 2 2 2 2 1 1 3 2 2 2 2 2 2 3 2 2 2 1
[75] 2 2 2 2 2 2 0 1 2 2 1 2 2 2 1 1 0 0 0 0 0 0 0 0 0
```

Level III: Translate the landscape reference scenarios to LANDIS-II.

Step 5	==>	Operationalize the LANDIS-II model for each of the HIP subclass reference systems
	==>	"Operationalize" means to parameterize, calibrate (validate), and define the simulation settings for each model

The proposed algorithm

Level IV: Build a landscape reference scenario codebook.



$= \{\text{aabbcdddd}\}$

MC Escher

Level IV: Build a landscape reference scenario codebook.

Step 6 ==> Perform multiple simulation runs (min r=3) for each of the HIP subclasses

==> Each dataset is a time series ($t = 20$) of the n-dimensional REVV; $\sim$ 30 data sets generated

==> Transform all raw output data (i.e., all REVVs) to KMVVs keeping time stamps intact

Level IV: Build a landscape reference scenario codebook.

Step 7 ==> Define a set of unique landscape morphs (states) for each landscape reference scenario

==> Morphs are derived from the collection of KMVVs generated by LANDIS-II using a feature extraction method

Many methods for identifying unique landscape morphs.
Choose one or more.

Hidden Markov Model?
Multi-section Vector Quantizer?
Support Vector Machine?
k-means Non-hierarchical Clustering Algorithm?
Principal Components Analysis?
Hierarchical Neural Network?
Logistic Regression?
Discriminant Analysis?

And many measures can be used by the methods!
Choose one or more.

Information measures
Distance measures
Dependence measures
Consistency measures
Accuracy measures

Level IV: Build a landscape reference scenario codebook.

Step 8a ==> Codify the morphs generated by each landscape reference scenario by giving each one a unique character symbol

==> Compile the symbols, their prototype (reproduction) KMVVs, and their probability distributions into a codebook

==> A "codebook" is a collection of symbols, prototype (reproduction) KMVVs, and probability distributions

Level IV: Build a landscape reference scenario codebook.

Step 8b ==> Combine the five codebooks generated by HIP subclass1 into one codebook, Alphabet1

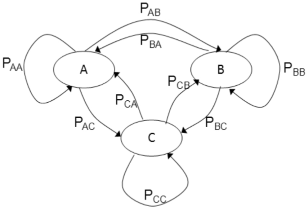
==> Combine the five codebooks generated by HIP subclass2 into one codebook, Alphabet2

Level IV: Build a landscape reference scenario codebook.

Step 8c ==> Define the complete landscape reference scenario codebook as the collection of Alphabet1 plus Alphabet2

The proposed algorithm

Level V: Build m-order Markov chains (bigram models) for the reference landscape scenarios.



$$\mathbf{P}=\begin{bmatrix} p_{AA} & p_{AB} & p_{AC} \\ p_{BA} & p_{BB} & p_{BC} \\ p_{CA} & p_{CB} & p_{CC} \end{bmatrix}$$

Time series data for 3 morphs on the landscape = AABBBABCBBBCBAA...

Level V: Build m-order Markov chains (bigram models) for the reference landscape scenarios.

Step 9 $==>$ Apply the landscape reference scenario codebook to the original KMVV data set to obtain landscape morph time series data (sequences of symbols) for each reference scenario

$==>$ This process will generate 10 sets of symbolic data

Level V: Build m-order Markov chains (bigram models) for the reference landscape scenarios.

Step 10a $==>$ Extract empirical frequencies from the time series data and use them to derive 1-step conditional probabilities

$$Logical\ product\ rule\ (Bayes)$$

$$\begin{aligned} p(AB) &= p(A|B)p(B) = p(B|A)p(A) \\ p(A|B) &= {}^{p(AB)}/_{p(B)} \\ p(B|A) &= {}^{p(AB)}/_{p(A)} \end{aligned}$$

Level V: Build m-order Markov chains (bigram models) for the reference landscape scenarios.

Step 10b $==>$ Define a bigram model from the conditional probabilities

$$Transition\ (stochastic)\ matrix$$

$$\begin{aligned} p(A|A) &= .25 \\ p(B|A) &= .75 \\ p(A|B) &= .10 \\ p(B|B) &= .90 \end{aligned} \quad = \begin{bmatrix} .25_{A|A} & .75_{B|A} \\ .10_{A|B} & .90_{B|B} \end{bmatrix}$$

The proposed algorithm

Level VI: Define agro-ecological sustainability for a given landscape.

$$
\begin{aligned}
&Landscape\ morphs \\
&\qquad A = morph_1 \\
&\qquad B = morph_2 \\
&\qquad C = morph_3 \\
&\qquad ... \\
&A\ landscape\ sustainability\ shift\ space \\
&\left[A \geq pr_1,\ B \leq pr_2,\ pr_3 \leq C \leq pr_4,\ ...\right]
\end{aligned}
$$

Level VI: Define agro-ecological sustainability for a given landscape.

Step 11 $==>$ Evaluate the ecological implications of the correlations evident in the transition (stochastic) matrices

$$\begin{bmatrix} .50_{A|A} & .50_{B|A} \\ .50_{A|B} & .50_{B|B} \end{bmatrix} vs. \begin{bmatrix} .75_{A|A} & .25_{B|A} \\ .95_{A|B} & .05_{B|B} \end{bmatrix}$$

$$\begin{bmatrix} 1_{A|A} & 0_{B|A} \\ 0_{A|B} & 1_{B|B} \end{bmatrix} vs. \begin{bmatrix} 0_{A|A} & 1_{B|A} \\ 1_{A|B} & 0_{B|B} \end{bmatrix}$$

Level VI: Define agro-ecological sustainability for a given landscape.

Step 12 ==> Use ecological and biological principles to evaluate agro-ecological sustainability by examining the fluxes in the components of the KMVV relative to the sequence of morphs appearing over time

Are the arbuscular mycorrhizal fungi increasing? Decreasing? Holding steady?

Level VI: Define agro-ecological sustainability for a given landscape.

Step 13 $==>$ Ideally, we would be able to evaluate agro-ecological sustainability on a given landscape by defining a shift space where the "forbidden blocks" are thresholds in the stationary distribution of landscape morphs

$$\left[A \geq pr_1,\ B \leq pr_2,\ pr_3 \leq C \leq pr_4,\ ...\right]$$

The proposed algorithm

Level VII: Explore the sustainability vocabulary of extant and novel landscape designs.

a food forest

By golly, it just might work!

Some experimental results

Some experimental results

A simple experiment:

(1) define a symbol generator as a 2-state Markov switching model

(2) use the switching model to simulate a time series data set (t=1000)

(3) use the empirical frequencies in the time series data to extract the conditional probabilities

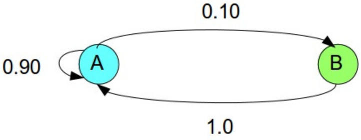
(4) use the empirically-derived conditional probabilities to define the empirical, 1-step Markov chain

(5) compare the empirically-derived Markov chain to the Markov chain in the symbol generator

(6) repeat (1)-(5) with another symbol generator model

(7) try a bootstrap version

Some experimental results



A simple, two-morph Markov chain used as a symbol sequence generator. The conditional transition probabilities are given as the edge set.

Simulated (time series) output from this generator:
ts.data = { A A A A B A A A A A B A A B A A A A A A A A A ... }

Some experimental results

```
G1<-matrix(c(.9,.1,1,0), 2, byrow=TRUE)
G1
     [,1] [,2]
[1,]  0.9  0.1
[2,]  1.0  0.0

eigen(G1)
$values
[1]  1.0 -0.1

library(expm)
tiks <- 15
a <- matrix(0, 4, 1)
for (i in 2:tiks){
 a <- cbind(a, c(G1 %^% i))
 if(identical(round(a[,i],5), round(a[,i-1],5)))
 {  break
  }
}
```

```
a
     [,1] [,2]  [,3]   [,4]    [,5]     [,6]      [,7]
[1,]    0 0.91 0.909 0.9091 0.90909 0.909091 0.9090909
[2,]    0 0.90 0.910 0.9090 0.90910 0.909090 0.9090910
[3,]    0 0.09 0.091 0.0909 0.09091 0.090909 0.0909091
[4,]    0 0.10 0.090 0.0910 0.09090 0.090910 0.0909090

limit.matrix1 <- round(matrix(a[,7], 2, byrow=FALSE),3)
limit.matrix1
      [,1]  [,2]
[1,] 0.909 0.091
[2,] 0.909 0.091

stat.dist1 <- limit.matrix1[1,]
stat.dist1
[1] 0.909 0.091
```

Some experimental results

```
library(distr)
 #Experiment1; Generator1; Dist1
exp1gen1d1 <- DiscreteDistribution (supp = c(1, 2),
prob = c(0.9, 0.1))
 #Experiment1; Generator1; Dist2
exp1gen1d2 <- DiscreteDistribution (supp = c(1, 2),
prob = c(1, 0))

set.seed(74)
tiks<-1000
Y_0<-1   #at t=0 ==> State 1
Y.G1.data <- NULL #empty vector
Y.G1.data[1] <- Y_0

for (i in 2:tiks) {
 if (Y.G1.data[i-1]=="1"){
  Y.G1.data[i] <- r(exp1gen1d1)(1)
 }else if (Y.G1.data[i-1]=="2"){
  Y.G1.data[i] <- r(exp1gen1d2)(1)
 }
}
```

```
Y.G1.data
  [1] 1 1 1 1 2 1 1 1 1 1 2 1 1 2 1 1 1 1 1 1 1 ...
 [39] 2 1 2 1 1 1 1 1 1 1 2 1 1 1 1 1 2 1 1 1 1 ...
 [77] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 ...
[115] 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 ...
..............................................
[837] 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 2 ...
[875] 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 ...
[913] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 ...
[951] 1 1 1 1 1 1 1 1 2 1 1 1 1 1 2 1 1 1 1 1 1 ...
[989] 1 1 1 2 1 1 1 1 1 2 1 1
```

Some experimental results

```
Y.G1.data.ts <- as.ts(Y.G1.data)
acf(Y.G1.data.ts, type = "correlation",
main="Y.G1.data")

p.a <- sum(Y.G1.data=="1")/length(Y.G1.data)
[1] 0.906

p.b <- sum(Y.G1.data=="2")/length(Y.G1.data)
[1] 0.094

library(gtools)
permutations(n=2,r=2,v=letters[1:2],
repeats.allowed=TRUE)
     [,1] [,2]
[1,] "a"  "a"
[2,] "a"  "b"
[3,] "b"  "a"
[4,] "b"  "b"
```

```
#automated counting window
count.bigrams <- NULL
for (i in 1:999) {
   #bi-gram {11} ==> "1"
if(identical(window(Y.G1.data.ts, i, i+1)[1:2], c(1,1))){
count[i]<-1
}else{
   #bi-gram {12} ==> "2"
if(identical(window(Y.G1.data.ts, i, i+1)[1:2], c(1,2))){
count[i]<-2
}else{
   #bi-gram {21} ==> "3"
if(identical(window(Y.G1.data.ts, i, i+1)[1:2], c(2,1))){
count[i]<-3
}else{
   #bi-gram {22} ==> "4"
if(identical(window(Y.G1.data.ts, i, i+1)[1:2], c(2,2))){
count[i]<-4
}
      }
    }
  }
}
```

Some experimental results

```
table(count.bigrams)
count.bigrams
  1   2   3
811  94  94

p.aa <- 811/1000
p.ab <- 94/1000
p.ba <- 94/1000
p.bb <- 0/1000
---- results: Pr(a|a) ----------------
p.a_a <- p.aa/p.a
[1] 0.8962472
---- results: Pr(a|b) -------------------
p.a_b <- p.ba/p.b
[1] 1
---- results: Pr(b|a) --------------------
p.b_a <- p.ab/p.a
[1] 0.1037528
---- results: Pr(b|b) --------------------
p.b_b <- p.bb/p.b
[1] 0
```

```
empirical.G1<-round(matrix(c(p.a_a, p.b_a, p.a_b, p.b_b),
2, byrow=TRUE), 3)

empirical.G1
       [,1]  [,2]
[1,] 0.896 0.104
[2,] 1.000 0.000

library(expm)
tiks <- 15
b <- matrix(0, 4, 1)
for (i in 2:tiks){
  b <- cbind(b, c(empirical.G1 %^% i))
  if(identical(round(b[,i],5), round(b[,i-1],5))) {
    break
  }
}

emp.stat.dist1
[1] 0.906 0.094
```

Some experimental results

```
library(entropy)
   #5 different hypothesized stationary prob dists
   #d3 = the empirically-derived stationary distr
   #d_true = the known stationary distr
d1<- c(.99,.01)
d2<- c(.95,.05)
d3<- c(0.906, 0.094)
d4<- c(.90,.10)
d5<-c(.85,.15)
d_true<- c(0.909, 0.091)

#as matrix
dM <- t(as.matrix(data.frame(d1,d2,d3,d4,d5)))
    [,1]  [,2]
d1 0.990 0.010
d2 0.950 0.050
d3 0.906 0.094
d4 0.900 0.100
d5 0.850 0.150
```

```
  #calc KL divergence for each distr
  #place results in a matrix, KLM
KLM <- matrix(0, 5, 1)
for (i in 1:5){
KLM[i,1] <- KL.empirical(d_true, dM[i,], unit="log2")
}

KLM
             [,1]
[1,] 1.779721e-01
[2,] 2.076308e-02
[3,] 7.696969e-05  #closest to known stationary distr
[4,] 6.673603e-04
[5,] 2.239388e-02
```

Some experimental results

```
#bootstrap for a (small) time series data set
(t=51)
Y.G1.data
 [1] 1 1 1 1 2 1 1 1 1 1 2 1 1 2 1 1 1 1 1 ...
 [39] 2 1 2 1 1 1 1 1 1 1 2 1 1 1 1 1 2 1 1 ...
 [77] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 ...
[115] 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 ...
..........................................
[951] 1 1 1 1 1 1 1 1 2 1 1 1 1 1 2 1 1 1 1 ...
[989] 1 1 1 2 1 1 1 1 1 2 1 1

sampled.output.G1 <-as.ts(Y.G1.data[250:300])
```

```
stat1 <- function(tsb){
count <- NULL
for (i in 1:50) {
if(identical(window(tsb, i, i+1)[1:2], c(1,1))){
  count[i]<-1
  }else{
if(identical(window(tsb, i, i+1)[1:2], c(1,2))){
  count[i]<-2
  }else{
if(identical(window(tsb, i, i+1)[1:2], c(2,1))){
  count[i]<-3
  }else{
if(identical(window(tsb, i, i+1)[1:2], c(2,2))){
  count[i]<-4
}
      }
    }
  }
}
a <- c(sum(tsb=="1"), sum(tsb=="2"))
b<-table(count)
c(a[1], a[2], b[1], b[2], b[3], b[4])
}
```

Some experimental results

```
library (boot)
set.seed <- 477
boot1<-tsboot(sampled.output.G1, stat1, 1000, 5,
"geom")

boot1$t[is.na(boot1$t)]<-0
bs.means <- colMeans(boot1$t[,1:6])
[1] 46.994  4.006 42.212  3.860  3.865  0.063

(bs.p.a <- bs.means[1]/50)
(bs.p.b <- bs.means[2]/50)

(bs.p.aa <- bs.means[3]/50)
(bs.p.ab <- bs.means[4]/50)
(bs.p.ba <- bs.means[5]/50)
(bs.p.bb <- bs.means[6]/50)

bootstrap.G1
         [,1]       [,2]
[1,] 0.9080521 0.09194791
[2,] 0.9745382 0.02546181
```

```
#convergence after 7 iterations
btstrp.stat.dist1
[1] 0.914 0.086

  #KL for 5 different hypothesized stat. prob. dists
  #[3,]=the bootstrap-derived stationary dist
            [,1]
[1,] 0.1779721139
[2,] 0.0207630769
[3,] 0.0002255110 #still the closest to the known distr
[4,] 0.0006673603
[5,] 0.0223938764

#compare KL values
KL for empirically-derived stationary distr
0.000077 = 77 x 10^-6

KL for bootstrap-derived stationary distr
0.000230 = 230 x 10^-6
```

Not bad! (KL $=$ 0 implies identical prob dist.s)

Lots to do!

Pitfalls and challenges

* defining state variables that capture multiple scale dynamics
* defining phase space partitions
* defining the landscape reference scenarios
* ag in LANDIS-II
* dangers of using Markov chains
* the ingression of novelty

Lots to do!

Next steps

* evaluate state variables
* evaluate ag in LANDIS-II
* run tests with 'dummy' landscapes

That's it folks!

Thanks.

Vincent's orchard