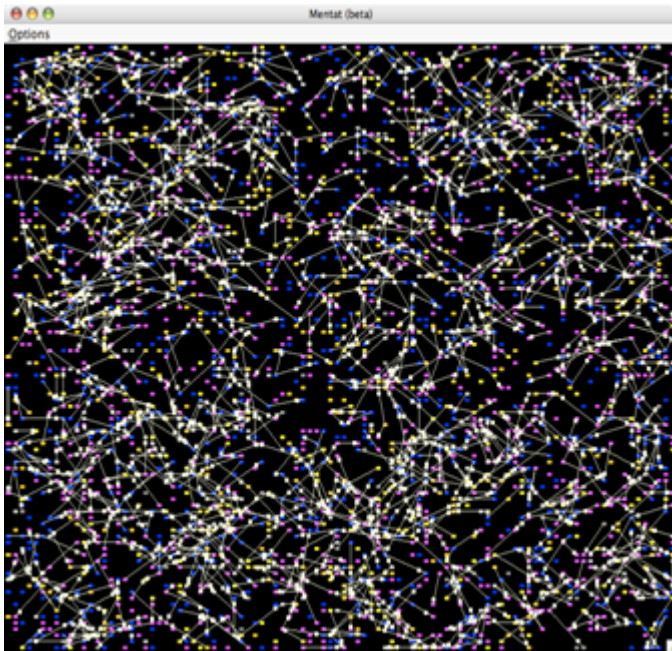




Injecting Data into Simulation: Can Agent-Based Modelling Learn from Microsimulation?



Samer Hassan

Juan Pavón

Nigel Gilbert



Universidad Complutense de Madrid



University of Surrey

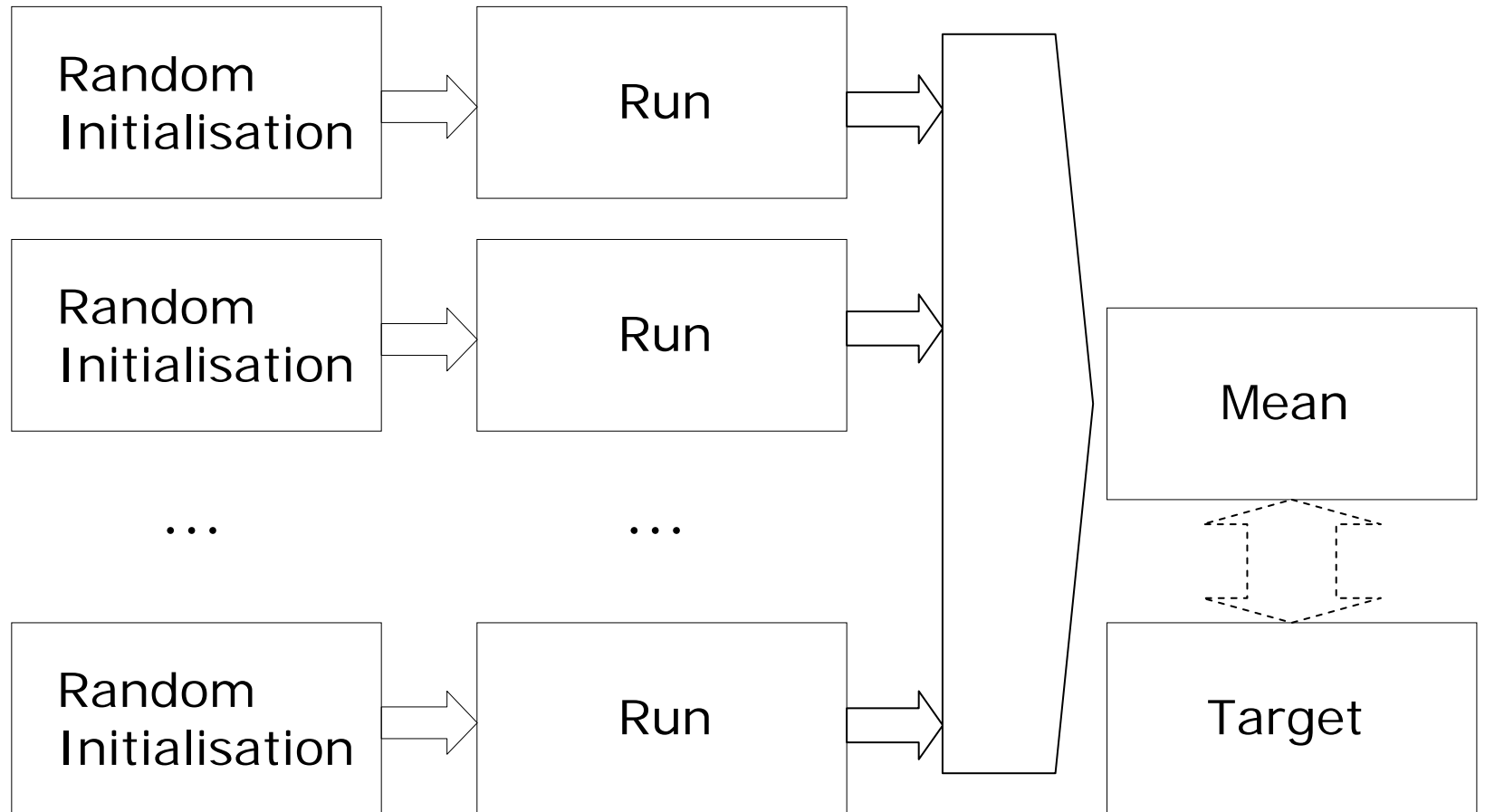
Contents

- Problems in Randomness
- A Method for Data-Driven ABM
- A Case Study: Mentat
- Concluding suggestions

Problems in Randomness

- Uniform Random Distribution
 - Common for initialisation
 - But also in
 - Distribution of objects in space
 - Determining unmeasured exogenous factors
 - Controlling agent behaviour

Problems in Randomness



Problems in Randomness

- There is always a chance of non-matching
- What if the target behaviour is an outlier?

Problems in Randomness

- Basing initial conditions on empirical data
- Moving ABM in the direction of the Target

An example: Eurovision song contest

- Hypothesis: “over a sufficiently long period of time the results of Eurovision would approximate to random”
- Random initial conditions & random voting schema should approach the real situation...
- ...but they don't

An example: Eurovision song contest

- Introducing empirical data approaches the real scenario:
 - Distance between countries
 - Measuring similarity of cultures

Contents

- Problems in Randomness
- A Method for Data-Driven ABM
- A Case Study: Mentat
- Concluding suggestions

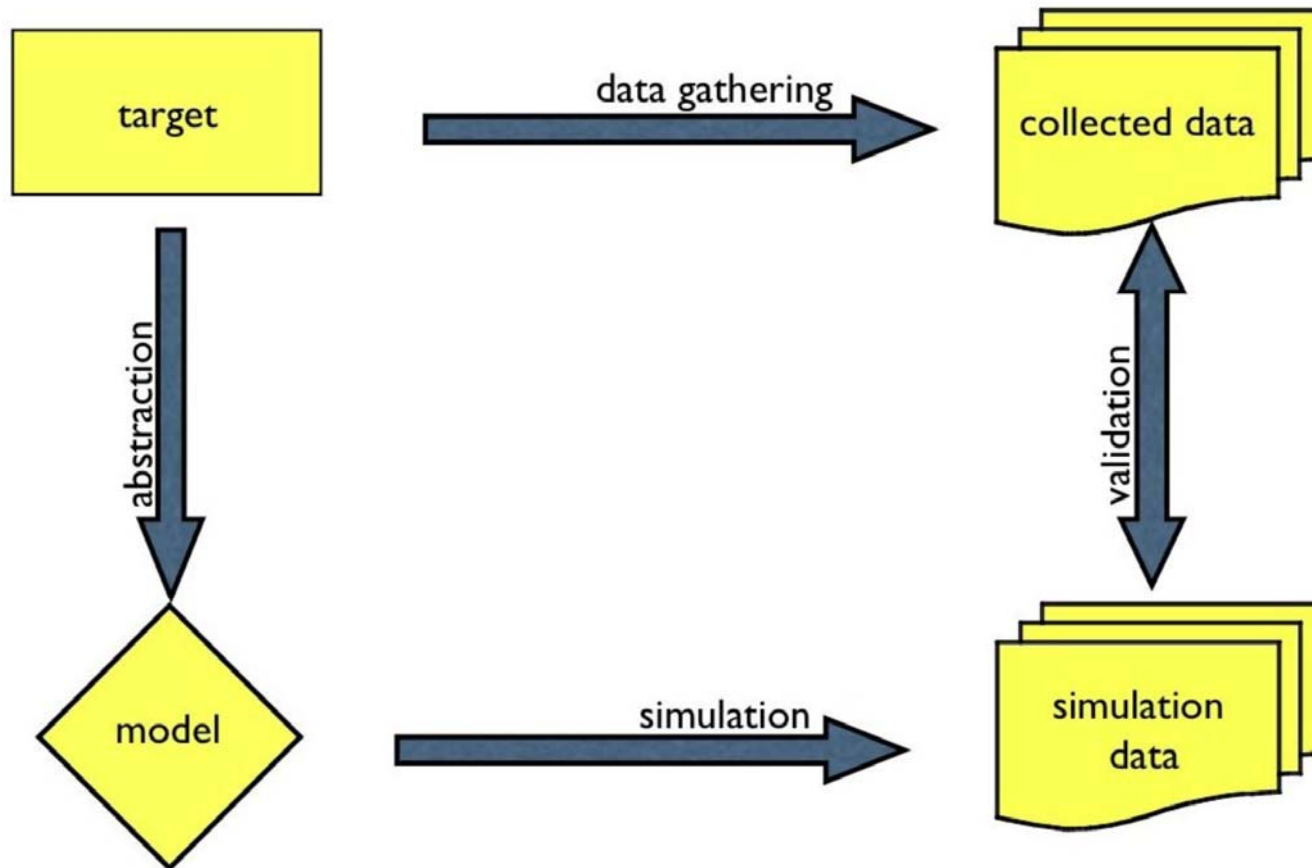
Microsimulation

- Approaching to Microsimulation
 - Surveys / Census → initialisation
 - Equations / Probability rules → behaviour
- Difficulties of Microsimulation
 - Requires plenty of quantitative data
 - Unable to model interactions

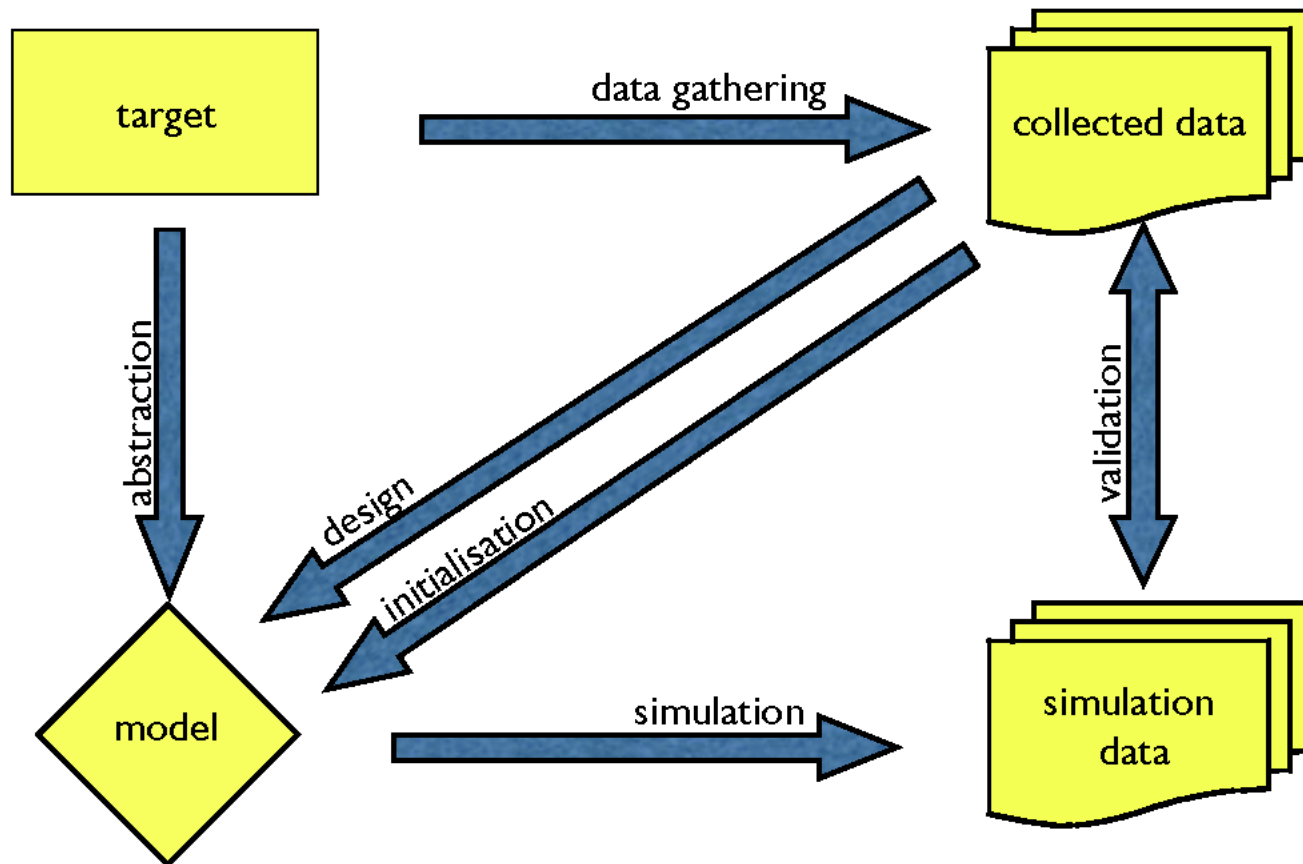
A Method for Data-Driven ABM

- Learning from Microsimulation:
 - Minimizing random initialisation
 - Basing the simulation in representative survey samples
 - Explicit rules need plenty of data
 - Using probability equations to determine changes in the values of agent parameters
 - Injecting more data into ABM
 - From other sources (e.g. qualitative)
 - In other stages (e.g. design)

Classical Logic of Simulation



Proposal for Data-Driven ABM



A Method for Data-Driven ABM

■ Difficulties

- When the ABM is too abstract
 - Empirical data cannot be obtained
- Requires detailed data from individuals
 - Suitable surveys? Unobservable?
 - Need of individual history? (panel studies)
- Requires dynamic information
 - Difficult to obtain: networks, micro-interaction
- Complicating not always implies benefits
- Loss of generality? Discussed

Contents

- Problems in Randomness
- A Method for Data-Driven ABM
- **A Case Study: Mentat**
- Concluding suggestions

A Case Study: Mentat

- Aim: simulate the process of change in moral values
 - in a period
 - in a society
- Plenty of factors involved
 - Now focusing on demography

Mentat: architecture

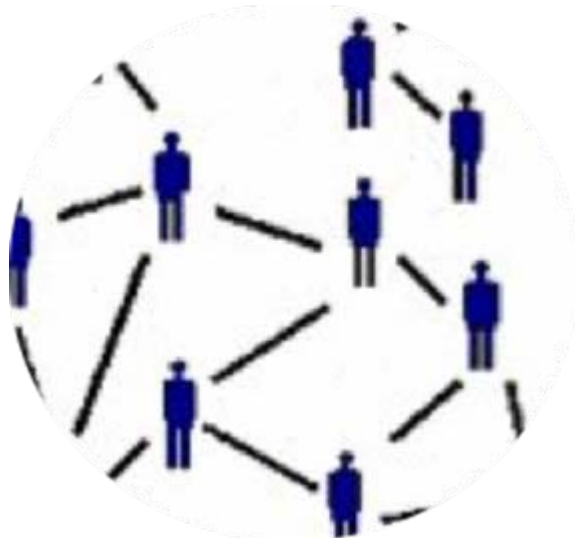
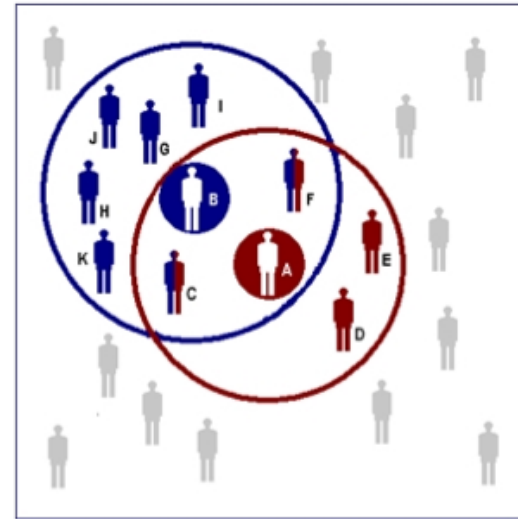
■ Agent:

- Mental State attributes
- Life cycle patterns
- Demographic micro-evolution:
 - Couples
 - Reproduction
 - Inheritance

Mentat: architecture

■ World:

- 3000 agents
- Grid 100x100
- Demographic model
- 8 indep. parameters



■ Network:

- Communication with Moore Neighbourhood
- Friends network
- Family network

A Case Study: Mentat

- Does the empirical initialisation substantially change the output in a pre-designed ABM?
- Random approach
 - No effort for additional data
 - Average behaviour
- Data-driven approach
 - Newly collected data is useful
 - Empirically based evolution

A Case Study: Mentat

- Two ABM:
 - Same design and micro-behaviour

 - Different initialisation
 - Mentat-RND: Random age
 - Mentat-DAT: Empirically based age

 - Same validation
 - Against newly collected data, not used in initialisation

A Case Study: Mentat

Comparison of outputs

	EVS/Census*			Mentat-RND			Mentat-DAT		
	1980	1990	1999	1980	1990	1999	1980	1990	1999
% 65+ years	16*	18*	21*	19	24	29	15	19	24
% Single	28	29	29	-	45	37	-	42	35
% Population Growth	-	-	+8%*	-	-	+10.1%	-	-	+7.2%

* Source: Spanish Population Census for the years 1981, 1991 and 2001

Contents

- Problems in Randomness
- A Method for Data-Driven ABM
- A Case Study: Mentat
- **Concluding suggestions**

Concluding suggestions

- Explore the problem background: availability of data?
- Compare different sources of data to give a stronger foundation to the model
- The most valuable data are those that provide repeated measurements
- Design ABM with an output directly comparable with empirical data
- Simulate the past and validate with the present

Thanks for your attention!

Samer Hassan

samer@fdi.ucm.es



University of Surrey



Universidad Complutense de Madrid

