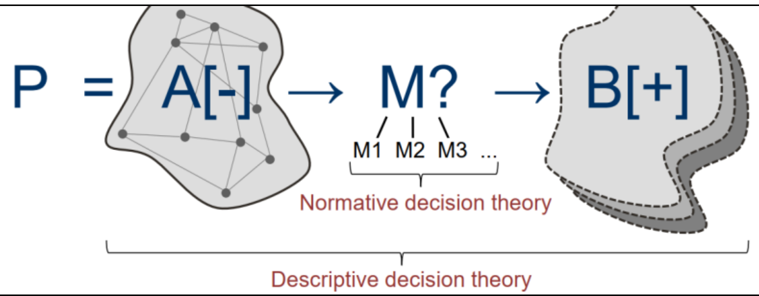
MANAGEMENT OF COMPLEX PROCESSES

0. Introduction

System	Set of **Elements** which are in a **relation** which a certain **purpose**. Biggest impact results, if you change the purpose, then the relation and finally the elements. e.g. Digestive (Verdauung) system.
Complex system	Complex systems are systems that **evade** (vermeiden) **simplification** and remain **multi-layered**. e.g. brain, human beings, internet, financial markets
properties	large, non-linear interactions, delayed effects, pos/neg feedback, network, open, universal, dynamic, robust, creative, innovative, unpredictable, differentiated sensitivity, not monitorable, ...
Complex vs Complicated	**complex: difficult to understand**, retroactive and side effects, delays and non-linearities (e.g. company) **complicated: large number of individual components**. Simplify by reduction (e.g. car, software)
Challenges	Analysis, design and management of complex systems (understand relationship, recognize variables) Decision-making in the context of complex systems (operative)
Models	Support the understanding of reality by simplification. It is always wrong, but still helpful. e.g. architecture, anatomical, physical, atomic, map, mental models.
Simulation	means using a model to **determine how things are going to act** in a certain situation (in all likelihood). is a **virtual experiment**. Which should answer the same questions as a corresponding real experiment,
properties	under controlled conditions (lab), more quickly, more cost-effectively, more resourcefully, in a non-dangerous way, within the framework of the model validity!
Decision Theory Overview	 Problem P, Situation A (worse) Model M (1..3...) Situation B (better, multiple)

1. Descriptive Decision Theory

Rational decision	The decision process is targeted throughout (durchwegs) and is consistently orientated towards targets. The ideas used in the decision process are based on the most objective information possible. The decision process follows a systematic procedure and uses clear methodological rules, which can be understood by those not involved.
Intuitive decision	one is usually conscious of only a small section of what one knows.
Mental model	is a simplified **mental depiction** (Darstellung) of the **subjective perception** (Wahrnehmung) **of the world** to • handle the large quantity of information taken in by the sensory organs • be able to **make decisions quickly** and process informations more efficiently. e.g. setting shower water, crossing the road, tying shoelaces (Schuhbänder), ...
properties	are **subconscious**, **gained by experience**, are **adapted to the specific situation immediately** are **important for survival** (information processing, efficiency increase, to interpret complex issues) are **generic terms** (formed by learning). They can lead to premature (voreiligen) conclusions are **implicit**. This is want makes them so dangerouse. -> Take a careful look.
Radical constructivism	The "reality" which surrounds us is the result of our ideas. Knowledge, insights, ideas are constructed by humans. e.g. The story of the hammer. Number selection game. Bavelas experiment (healty cell detection).
Correlation vs Causality	**Correlation**: Mathematical dependency between two factors, can be statistically proven! **Causality**: Cause-effect relationship exists between two factors, cannot be proven!

Problems of mental models	**Superstitiousness** (Aberglaube)	Ticket machine on the bus (coin robber, ticket machine)	Distorted perception of reality
	Prejudices (Vorurteile)	Rosenhan psychiatric clinic experiment (sick/healty people)	
	Generate their own momentum	Selective search after confirmation	
	Conservatism	Mental models develop based on experience	
	False security	If they are not perceived, they are not questioned (goats problem)	

Defuse problems	Make mental models plausible 1. become familiar with your own mental model, make mental models explicit 2. critical thinking 3. exchanging information about mental models Mental models cannot be verified!
Falsifiability	"A statement must be falsifiable in order for it to be scientific." e.g. all swans are white. Existence of several white swans does not prove -> search for one black one

2. Normative Decision Theory

Goal	The main area of activity of decision theory is not conflicting targets but rather **uncertainty**.
Environmental conditions	Factors which influences an alternative decision over which you don't have any influence.
State space	The space of all possible decision results. In reality this space is often not fully recorded.
Decision	involving **risk**: If occurence likelihoods can be assigned to environmental conditions. involving **uncertainty**: If no occurrence likelihoods can be assigned to environmental conditions. rational counterpart: environmental conditions that are not determined by coincidence
Game theory	A special branch of decision theory. e.g. Prisoner's Dilemma: **Collectively** it is of advantage for both to remain silent. (2 years each if silent or 4 years by confessing) **Individually** it appears more advantageous for both to give a statement (1 year instead of 6 years).
Dominance	If it is to be preferred over another alternative in all cases.
Absolute	The worst value of the dominating alternative is better than the best result of the dominating alternative.
Circumstantial	One alternative in every circumstance is better than or equally good as the other alternative.
Targets	**Complementary targets**: By pursuing (verfolgen) one target, the other target is also optimally achieved. **Neutral targets**: Pursuing one target does not influence the other target. **Competing targets**: Realising one target impacts on reaching the other target.

Utility analysis

Criterion	Weighting	A			B		
delivery time	0.2	immediately	6	1.2	25 weeks	4	0.8
support	0.2	sufficient	4	0.8	comprehensive	5	1
design	0.1	average	4	0.4	moder, fit company	5	1
price	0.4	very good	6	2.4	rather expensive	4	1.6
experience	0.1	none	1	0.1	very extensive	6	0.6
total	**1.0**			**4.9**			**4.6**

sum of factors = 1
scale: 1-6 or 0.0-1.0

if similar	more criteria, changing the weights
improve result	make survey (include more people), quantify, make hard results

target weighting utility function (linear)

	Revenue	factor	Profit	factor	Result = Revenue * factor + Profit * factor
a1	800'000	0.25	7'000	0.75	205'250
a2	600'000	0.25	8'000	0.75	156'000

standardise

$$standardised\ value = \frac{value\ to\ be\ standardised - minimum\ value}{maximum\ value - minimum\ value}$$

Decision involving uncertainty

no racionality at all

	z_1	z_2	z_3	z_4	min	max	$\phi(a_i)$ e.g. 0.4	$Laplace$
a_1	60	30	50	60	30	60	$0.4*60+0.6*30=42$	$\frac{60+30+50+60}{4}=50$
a_2	10	10	10	140	10	140	$0.4*140+0.6*10=62$	$\frac{10+10+10+140}{4}=42.5$
a_3	-10	100	120	130	-10	130	$0.4*130+0.6*(-10)=46$	$\frac{-10+100+120+130}{4}=85$

MaxiMin rule pessimistic: Choose the one with the **highest minimum**. Not according to reality. Only 1 value is considered.

MaxiMax rule optimistic: Choose the one with the **highest maximum**. Extremely not according to reality. Only 1 value is considered.

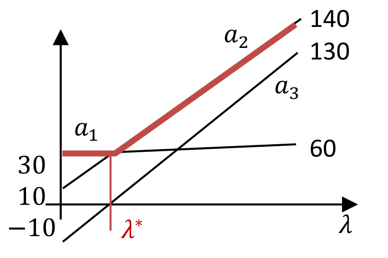
Hurwicz rule

$$\phi(a_i) = \lambda * Max_j(e_{ij}) + (1-\lambda) * Min_j(e_{ij})$$

$$(\text{low risk})\ 0 \le \lambda \le 1 (\text{high risk})$$

2 values are considered.

$$\lambda * 60 + (1-\lambda) * 30 = \lambda * 140 + (1-\lambda) * 10$$
$$\lambda * 30 + 30 = \lambda * 130 + 10$$
$$20 = \lambda 100$$
$$\lambda^* = 0.2$$
$$0 \le \lambda \le 0.2 \rightarrow a_1\ \ 0.2 \le \lambda \le 1 \rightarrow a_2$$



Laplace criterion: take the average. only rational for risk neutral decision makers

Savage-Niehans rule "lowest regret"

		z_1	z_2	z_3	z_4	3. max-> choose min
1. max per column	max	60	100	120	140	
2. regret per alternative	a_1	60-60=0	100-30=70	120-50=70	140-60=80	80
	a_2	60-10=50	100-10=90	120-10=110	140-140=0	110
	a_3	60-(-10)=70	100-100=0	120-0	140-130=10	70

Not all information is used. Pessimistic approach.

Decision involving risk

e.g. profit maximisation of two petrol stations		Circumstances / Zustand z1 (no bipass) $p_1 = 0.7$	 z2 (bypass) $p_2 = 0.3$	μ (expected value) / Bayes rule
Alternative	**a1**	$125'000$	$125'000$	$0.7 * 125'000 + 0.3 * 125'000 = 125'000$
	a2	$150'000$	$80'000$	$0.7 * 150'000 + 0.3 * 80'000 = 129'000$

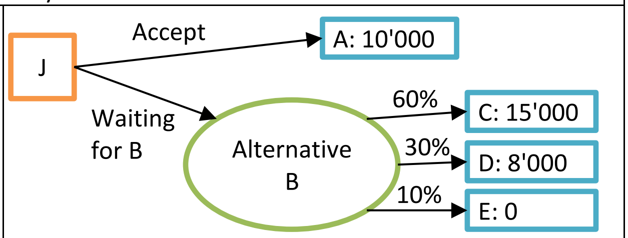
risk aversion (Abneigung): prefer a lower expected value if this provides more security
risk neutral: only see expected value -> rational
Criticism: **venturesome**: avoid a higher expected value in favour of a wide range of varying results

St.Petersburg paradox: In this game, a coin is flipped and the game is ended as soon as "heads" appears. The prize starts at 2 and doubles every time "tails" appears. How high would be your stake be?

Decision tree

are ordered structured trees which can be used to show decision rules. The graphical depiction as a tree diagram illustrates decisions which follow one another in a hierarchical order. Elements:

- root node
- inner nodes
- leaf nodes (answer)



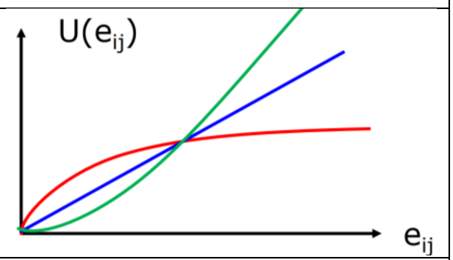
EMV: "Expected monetary value"
At 11'400 waiting for B has a higher expected value.

properties:
The time runs from left to right.
Event nodes are display as circles. Decision nodes as squares.
Write down all possible events/decisions.

Bernoulli principle

Uses a risk utility function (RNF). Select the alternative for which the greates expected utility results.
e.g. Whould you likely have 1 milion money for sure or 2 million for 50%?

blue	risk neutral	straight curve
red	risk aware / risk averse	concav (right bended)
green	adventuresome / risk seeking	convex (left bended)



Example

Result matrix:

e_{ij}	z_1 0.5	z_2 0.2	z_3 0.3
a_1	30	20	20
a_2	140	−40	−30
a_3	40	10	60

Utility or decision matrix:

$U(e_{ij}) = 300 * e_{ij} - e_{ij}^2$	z_1 0.5	z_2 0.2	z_3 0.3	$E\left(U(e_{ij})\right)$
a_1	$8'100$	$5'600$	$5'600$	$= 6'850$
a_2	$22'400$	$-13'600$	$-9'900$	$= 5'510$
a_3	$10'400$	$2'900$	$14'400$	$= 10'100$

Decision rule $\Psi(\mu, P_0) = \mu - 10 * P_0$
$\mu: expected\ value$
$P_0: probability\ of\ ruin(e.g. < 21)$

	μ	P_0	$\Psi(\mu, P_0)$
a_1	25	0.5	20
a_2	53	0.5	48
a_3	40	0.2	38

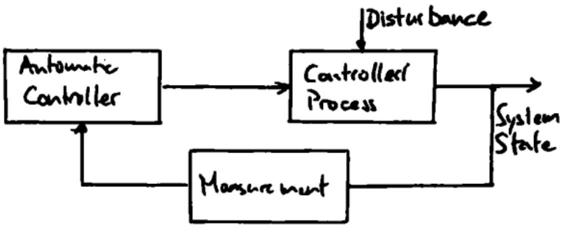
Utility function

$$U = 2 * \mu - 10 * \begin{cases} 1\ if\ \mu < 21 \\ 0\ if\ \mu \geq 21 \end{cases}$$

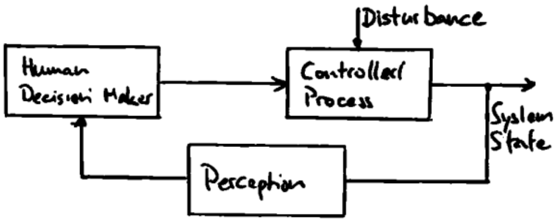
3. System Thinking and System Dynamics

Overview

Automatic Control System



System controlled by Humans

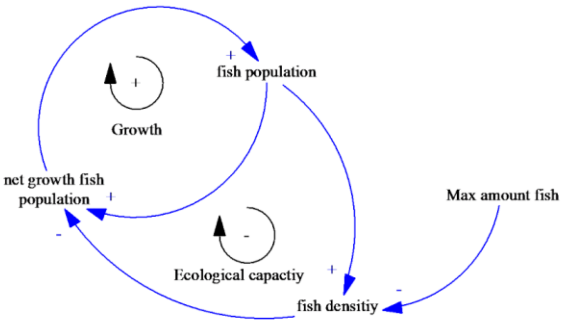


Humans are not designed, they are evolving.

Example: Fish population

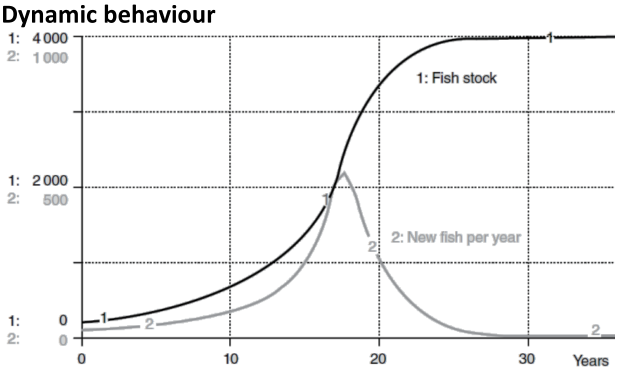
There is only a small amount of fish in the beginning.
The available amount of food serves only a certain amount of fish.

CLD = Cause and Loop Diagram



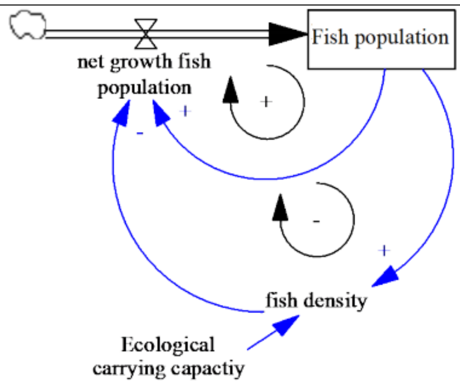
Elements: Variables and Causal Links

Dynamic behaviour



S-shaped growth of the fish population

Stock-and-Flow-Diagram



Elements

Stock	F	Fish population
Flow	$\frac{dF}{dt}$	net-growth fish population $\frac{dF}{dt} = f(F, d) = f(F, d(F, CC))$
Auxiliary Variable	d	Fish density $d = d(F, CC) = \frac{F}{CC}$
Model Parameter	CC	Ecological carrying capacity

Polarity

Positive Polarity

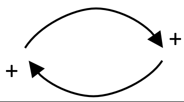
if A increase, B increase
if A decrease, B decrease

Negative Polarity

if A increase, B decrease
if A decrease, B increase

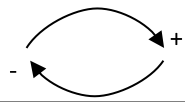
Feedback Loop

Reinforcing Loop



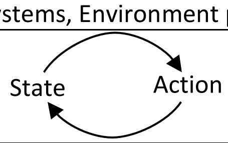
even number of minus, reinforcement increases the initial effect

Balancing Loop



odd number of minus, balancing of the initial effect

Basic concepts

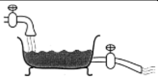
History	"System Dynamics": Modelling and simulation method, developed by Jay W. Forrester (MIT) 1972
Application	Industry/Production, Management control (Strategy planning), Macroeconomics, Social systems, Health systems, Environment planning (Analysis of Human-environment-Systems)
Feedback thinking	
Endogenous perspective	We derive the essential dynamics from the mechanisms within the system boundaries. "almost nothing is exogeneous" - John Sterman endogen -> internal growth/cause exogeneous -> external cause
Causal thinking	Structure (CLD) drives behavior (dynamic behavior diagram)
Goals	**Understanding** the interactions in a complex system that are conspiring to create a problem and understanding the structure and dynamic implications of policy changes intended to improve the systems behaviour (Richardson 1991), not forecasting.

	Not to build the model of the system, but rather **to get a group engaged in** build a system dynamics model of a problem in order to see what extent this process might be helpful to increase problem understanding and to devise courses of action to which team members feel committed (Vennix 1996).
Shared mental models	The actors within the system develop shared mental models and decision making rules. They change the structure (CLD, e.g: by limiting fishery) and change the behavior (dynamic behavior diagram)
3 Levels of formalisation	1. Causal Loop Diagrams (Communication tool for a shared mental model and general discussion) 2. Stock and Flow Diagrams (Precise visualization of stocks and flows) 3. Stock and Flow Diagrams with equations (Enables quantitative simulation)

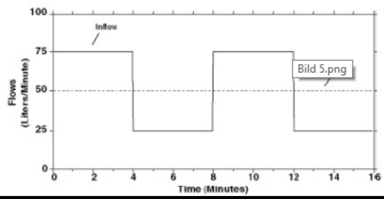
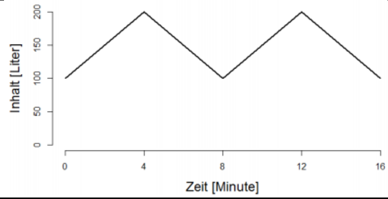
Accumulation (Delays and Feedbacks)

Key Message	Understanding accumulation is fundamental to understanding system behaviour. "A stock takes time to change, because flows take time to flow. That's a vital point, a key to understanding why system behave as they do". D.H. Meadows, 2008

Metaphor and Equation

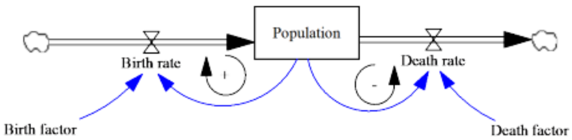
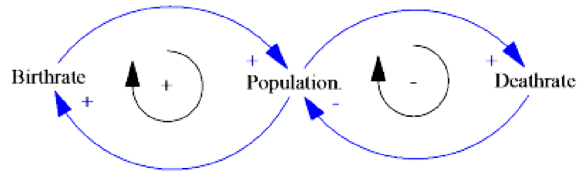
Stock-and-Flow-Diagram	Metaphor	Equation
		$Stock(t) = \int_{t_0}^{t} [Inflow(s) - Outflow(s)]ds + Stock(t_0)$ $\frac{d(Stock(t))}{dt} = Inflow(t) - Outflow(t)$

Bathtub dynamics

Problem	Given: Flow	Solution: Content
Draw the anticipated behaviour of the bathtub content over time the depicted diagram Initial content: 100l		

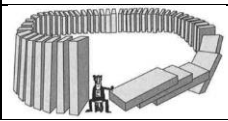
Causal Loop Diagrams

Compare

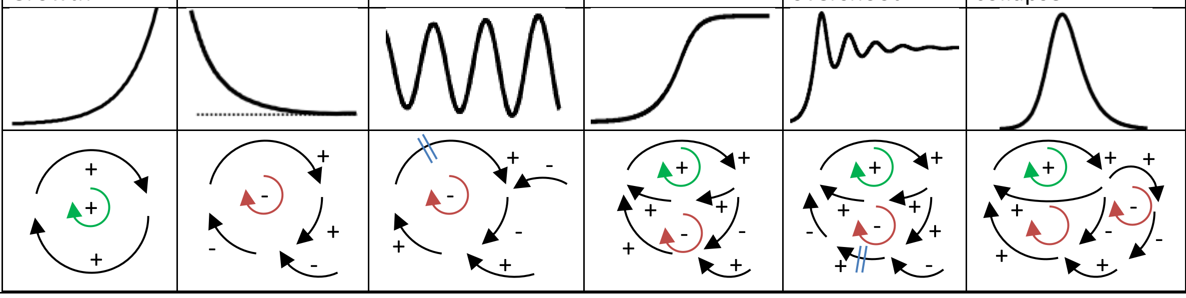
Stock-And-Flow-Diagram	CLD
	
Stocks are state values (the memory) of a system need an initial value only change via In- and Outflows Value does not change -> Inflow = Outflow Value increases -> Inflow > Outflow Value decreases -> Inflow < Outflow have units (such as liter, CHF, meter) **Flows** are changing rates of stocks have units (such as liter/h, CHF/year, m/sec)	Create an overview and understanding Uncover leverage points Reduce the complexity in finding potential side-effects and feedback from policy interventions Identify dependent variables & potentially conflicting goals Improve quality of discussion by precise presentation of the subject of matter

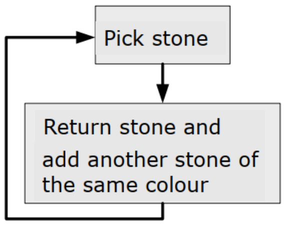
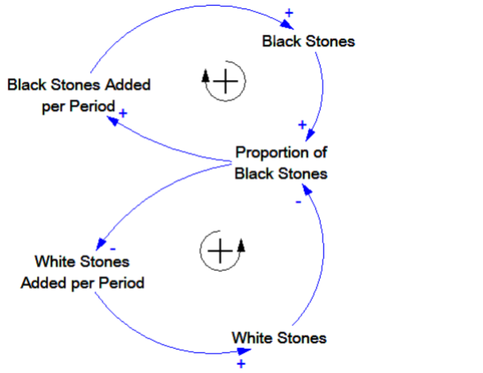
Causal relations

Instantaneous		Accumulative	
Increases A, then increases B	Increases A, then decrease B	A is added to B $B = \int_{t_0}^{t} A\,ds + A_0$	A is subtracted from B $B = \int_{t_0}^{t} -A\,ds + A_0$
Birthrate Population +	Birthrate Fish density -	+ Birthrate Population	Population Deathrate -

Best practices Variable Names	Select substantive/noun as variable name	Number of Installations
	Don't indicate the direction of variables via name	~~increasing~~ electricity generation
	The value of the variable needs to be able to increase or decrease. the name ideally describes a continuum	~~Municipality Fusion~~ Attractiveness of Municipality Fusion
Causal relations	Model causality not correlation - include only causality that convices you - finding relevant causal explanations is a creative mission and requires an exchange of information of relevant stakeholders	~~ice cream sales -> pool accidents~~ peak temp -> ice cream sales peak temp -> pool accidents
	Model causal relations with clear polarity. If no polarity can be indicated, the underlying structure needs extended analysis.	~~Income tax rate -> Tax revenue~~ income tax rate -> attractiveness of location -> tax subrate -> tax revenue
	No "Shopping List". Rule of thumb: Max 3 causes for 1 effect	~~training conditions -> sucess~~ ~~successfull scouting -> success~~ ~~teamspirit -> success~~ ~~quality of manager -> success~~
	Avoid redundancy leads to further insights	~~Market share -> Sales~~ ~~Salesforce -> Sales~~ Salesforce -> Market Share -> Sales
	Ask yourself with every decision rule: Are there unintended consequences?	open Tasks -> Extra time -> Task done ? Extra time -> Exhaustion
Loops	Mark the important delays with //	Extra time --//--> Exhaustion
	And close the loop	Exhaustion -> Tasks done
	Validate **all** feedback: Which story does the loop tell? - Bother about "invisible" feedback. Use mathematical polarity to validate the loop within the story.	
Model	Focus on the purpose - Which goal variable need to be modelled? Which stories are important to you and the target group? - Identify redundant relations If necessary further abstraction of variables	
Lesson Learnt	A CLD is more specific if connected to a "Reference mode of behavior" Define the variable as precisely as possible. The units are your friend. Don't model a system, model a problem. Model the policy (=Rules of decision making). By doing so, you will uncover the relevant feedback mechanism.	
Limits	Which loop dominates? When does the system collapse? How do loops interact with each other?	

Fundamental modes	Exponential Growth	Goal seeking	Oscillation	S-shaped growth	Growth with overshoot	Growth and collapse



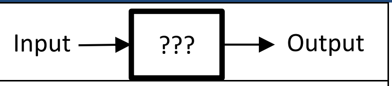
Polya Process	Path dependence and Lock-In effects. From György Pólya 1887-1985.	
Lock-in situations	- based upon decisions in the past - when changing the current situation is linked to high investment - and therefore often doesn't happen	QWERTY-Keyboard Bonus system, Coop to Lufthansa Gender roles, Fossil fuel powered civilization
Process	Initialising: n stones of each colour usually: n=1	 

Lessons learnt	A system might react with a different sensitivity to the same disruptions/interventions during different phases. During specific system conditions, a minimal intervention suffices to redirect the behaviour towards the desired direction. These conditions are often during the early stages of a development. During other conditions (often later, after a longer development) system change is often only possible with high investment.
Example	produce a part each 60 minutes, and 40 minutes is needed for one part $$p = \frac{\lambda}{\mu} = \frac{\frac{1}{60}}{\frac{1}{40}} = \frac{2}{3}$$ error rate of 20% $$p = \frac{\lambda}{\mu} * \sum_{n=0}^{\infty} 0.2^n = \frac{\lambda}{\mu} * \frac{1}{1-\epsilon} = \frac{2}{3} * \frac{1}{1-0.2} = \frac{5}{6}$$

4. Simulation

Modeling

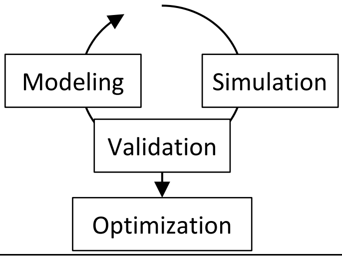
Modeling is Simulating!
Is the process of establishing the relationship of the input to the output.



Analysis Landscape

Original System	vs	**Model**
Physical Model	vs	**Computer Model**
Theorie-based Model (Whitebox)	vs	**Data-based Model** (Blackbox - Machine Learning)
Analytical Model (Formula)	vs	Simulation Model (more effort)

Workflow



Resources

Research
Data

Paradigms

Monte Carlo Simulation: broad class of algorithms that rely on repeated sampling to obtain numerical results.
System Dynamics (SD): approach for understanding the behavior of complex systems over time, using stocks, flows, feedback, loops, and time delays.
Discrete Event Simulation (DES): system state only changes at discrete points in time. key element: Queue
Multi-agent Simulation: system that consists of multiple entities with identical or different bhavior. They solve problems collectively. e.g. ants

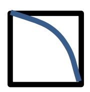
overview

	Agent Simulation	DE Simulation	System dynamics
What is being tracked?	Individual Objects	Individual Objects	Population
Process logic	Locally, inside objects	Central control unit	Central control unit
Time	Discrete	Discrete	Continuous
Typical Applications	Biological Processes; Human behavior, e.g. negotiations	Cleary described processes in areas Logistics and Production	Physical Processes, Political Processes, Socio- economical Processes

Monte Carlo Simulation

is a numerical method for statistical simulation.
It is based on using sequences of random numbers in order to explore the behavior of the system. Monte-Carlo simulation also provides an alternative for solving complex problems in probability and statistics.

calc Pi



$r^2 * \pi = A$
take 1000 rand points
count $x^2 + y^2 \leq 1$

Model

Scenario Generator: creates scenarios according to a prescribed probability distribution that match reality. Each scenario yields values of the target variable that can be obtained via a calculation scheme or formula.
The aggreator calculates the output from the target values associated with each simulated scenario.

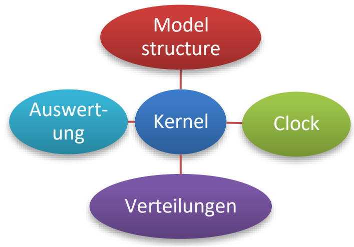
Example: Newsvendor problem

A newsvendor can decide freely every morning how much newspapers he wants to procure at the depot. Each newspaper costs 1.-.
He then goes to the central market square and offers newspapers for 1.50.
The daily demand for newspapers is stochastic and is uniformly distributed between 100 and 190. Unsold newspaper at the end are valueless.
How much newspapers must the newsvendor buy if he wants to maximize his profits?

Q: Order quantity
N_i: daily demands
G: Profit $= \min(N_i, Q) * p - Qc$
p: Salesprice
c: Purchase price
F: (Cumulative) Demand distribution
E: Expected profit $= \frac{1}{L}\sum_{i=1}^{L} G_i$

Discrete Event Simulation
A Simulation-model is a virtual representation of a real-world system that allows to gain insights in how the system works in reality. Key element in DES: **Queue**

properties
Presence of stochastic (random) variables
(Evolution over) Time pays an important role
Individual entities are being monitored (in constrast to SD)
State changes are the result of "events".
Events only take place a discrete points in time.
DES successfully applied in many different environments like shop floor control, spare parts logistics, pedestrian modeling, et cetera.



Queuing system
Stability of the system, Number of customers in the queue, average waiting times, 95% percentile of all waiting times.

Examples

System	Server	Job
Bank	Teller	Consultancy
Hospital	Doctors, Nurses, Beds	Consultancy / Treatment
Computer	CPU, I/O-Devices	Jobs
Manufacturing	Machines, Service, Engineers	Measure, drilling, packing, ...
Ambulance services	Ambulances, doctors	Transport / Treatment

traffic systems, telecommunication, logistics, production, et cetera.

Advantages
Cost-effective, fast, and safe field of experimentation
Allows animation and hence increases system understanding
Allows the analysis of complex systems at a high level of detail

Disadvantages
Requires significant (software) development time
Construction of a simulation model is relatively prone to error
Sound interpretation of results is challenging

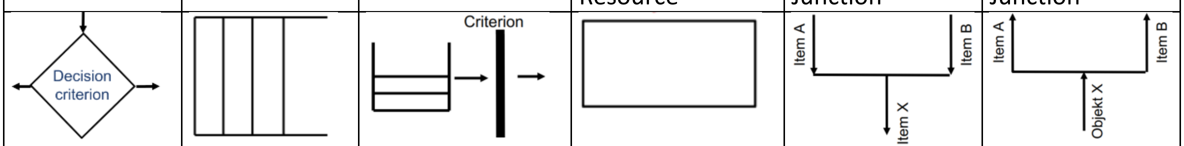
Elements
Items (or Entities) flow through the system (e.g. patients, products, spare parts)
- have attributes (arrival/processing time, priority)
Simulation clock holds the virtual time within the simulation model
System state $Z(t)$ gives a full description of the system at time t.
- current assignment of items (with their current attribute values) to the various queues and resource blocks.
Future Event List contains 0 or more (time, event) pairs. It contains those events that are known at the current time of the simulation clock. During a run, events are added and removed from this list continuously.
Each **Event** has an effect at the time of occurrence. The effect consists of a change to the system state and/or a change to the Future Event List. In the time between two consecutive events, nothing happens (the system state remains unchanged). Therefore the simulation clock can jump from an even to event.

Software Tools
Simulation models can be developed in any programming language.
Advantage: maximum flexibility
Disadvantage: requires significant implementation effort.
For most languages there exist stable (open-source) DES libraries — Simmer in R
Commercial software packages greatly simplify modeling and implementation by providing access to many DES "building blocks" — Arena, Simio, Matlab, AnyLogic

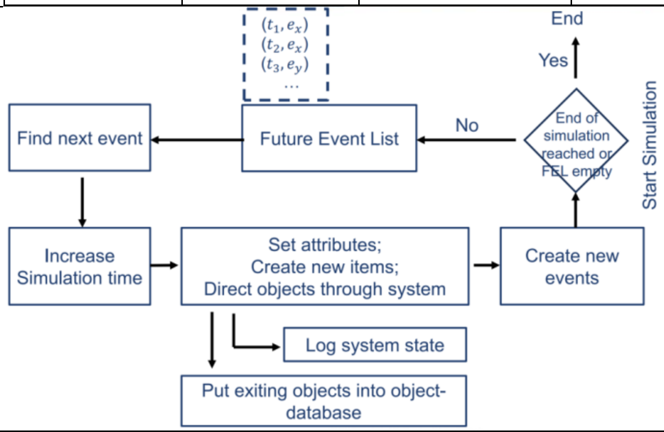
Flow &Event Diagrams
shows how items can move through the system and describe the impact of a given event on:
Attribute values of active items / System state / Future Event List

Elements

Switch	Queue	Gate	Activity / Resource	convergent Junction	divergent Junction



kernel



best practices

Avoid intersecting arrows

Add small explanations to arrows, resources and activites

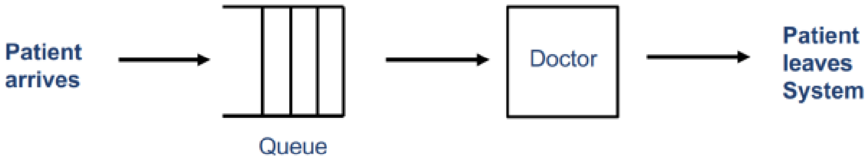
Consider building sub models if the number of components gets bigger than ~10

Example: Medical center

Problem description

We consider a medical center with 1 doctor. Each morning from 09:00 to 10.00 she is available for everybody who needs consultancy related to minor medical problems. Patients cannot make an appointment. They come and then wait until they are called in. Patients are served according FCFS (First-Come First Served). The patients arrival process and the consultancy times are both stochastic. We would like to know the average waiting time for various arrival rates and average service times.

Flow diagram



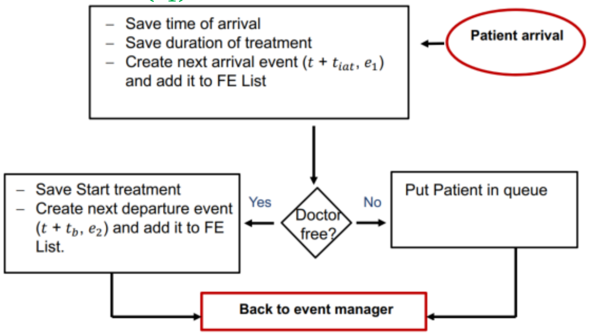
Object class model

Objects are the patients. In order to calculate the average waiting time, we must specify time of arrival and time when treatment starts for each patient. Further, we need the treatment duration in order to decide when a patient leaves the system. This means that we identify following attributes:

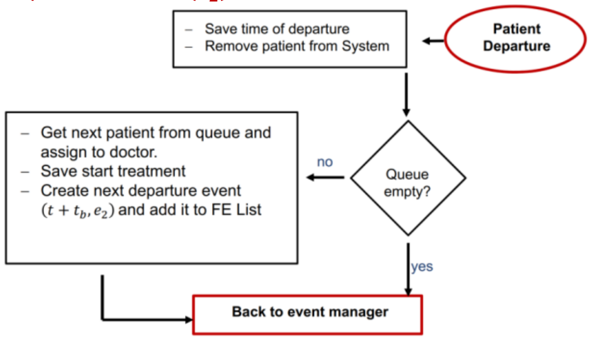
- Time of arrival
- Start treatment
- Treatment duration

Departure time is optional, because it can be obtained from the attributes "Start treatment" and "Treatment duration". In general, it is not a bad idea to save relevant information as attribute in the objects.

Arrival event (e_1)



Departure event (e_2)



Aggregator

$$E(W) \approx \frac{1}{L}\sum_{i=1}^{L}(x_i.start.treatment - x_i.end.treatment)$$

W: waiting time
L: number of simulated patients

Descrete Event Mechanics

Time	Event	Random numbers	FE-List	Objects	Doctor	Queue
0		IAT=5	$(5,e_1)$	$\emptyset$	$\emptyset$	$\emptyset$
5	e_1	PRT=3 IAT=4	~~$(5,e_1)$~~ $(8,e_2)$ $(9,e_1)$	$(P1;05;05;03)$	$P1$	$\emptyset$
8	e_2		~~$(8,e_2)$~~ $(9,e_1)$	~~$(P1;05;05;03)$~~	~~$P1$~~	$\emptyset$
9	e_1	PRT=3 IAT=2	~~$(9,e_1)$~~ $(12,e_2)$ $(11,e_1)$	$(P2;09;09;03)$	$P2$	$\emptyset$
11	e_1	PRT=4 IAT=1	~~$(11,e_1)$~~ $(12,e_2)$ $(12,e_1)$	$(P2;09;09;03)$ $(P3;11;-;04)$	$P2$	$P3$
12	e_2		~~$(12,e_2)$~~ $(12,e_1)$ $(16,e_2)$	~~$(P2;09;09;03)$~~ $(P3;11;12;04)$	$P3$	~~$P3$~~
12	e_1	PRT=4 IAT=1	~~$(12,e_1)$~~ $(16,e_2)$ $(13,e_1)$	$(P3;11;12;04)$ $(P4;12;-;04)$	$P3$	$P4$
13	e_1	PRT=4	~~$(13,e_1)$~~ $(16,e_e)$	$(P3;11;12;04)$ $(P4;12;-;04)$ $(P5;13;-;04)$	$P3$	$P4, P5$
…	…	…	…	…	…	…

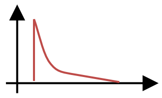
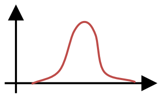
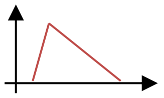
IAT = Interval Arrival Time
PRT = Patient Treatment
Objects: Patient ID (Id), Arrival time (Ta), Start treatment (Tst), Treatment duration (Ttd)
FE = Future Event

Stochastic processes

In statistics, uncertainties are modelled via **random variables**.
A serie of numbers is random if it does not contain any patterns.

types

Theoretical distribution
+ compact formulations
+ easy to adjust
+ wide range of returned numbers
- limited number of params
-> nor reality

Uniform	Exponential	Normal	Triangular
			
Interval [a,b] $a + m * x$	$\lambda = 0.01$ $-\lambda * \ln(1-x)$		min,max most likely

Empirical distribution
+ based on real data
+ any shape is possible
- not possible to generate new data
- cannot be described compact
- high memory footprint

Table with values

1	2	3	4	5	6
1023	1564	0	248	775	645

challenges

Picking the right distribution type
Estimating parameters values in selected probability distribution.
Getting data can take a long time
Data quality is often low and important documentation is missing

common mistakes

not enough data available
no data filtering
seasonal patterns ignored (summer/winter, mon-fri-sat-sun)
use of non-representative historical data
wrong interpretation

advice

-> always plot your data

Random-generator

Requirements
Independent (uncorrelated to previous realization)
Uniformly distributed
High resolution (no gaps)
No cycles
Fast and memory friendly creation
Reproducibility
Remark
Real random numbers cannot be reproduced!
But we might need reproducible random sequences, so called pseudo-random numbers.
Properties
Generation starts with some initial value (seed). This determines the entire series of random numbers.

Linear Congruency Method

$$x_{i+1} = (ax_i + c) \bmod m$$
$$u_i = \frac{x_i}{(m-1)} \sim U(0,1)$$

x_{i+1}	new value
a, m	very big prime numbers
c	parameter

$x_0(Seed) = 37, a = 19, c = 11, m = 833$

	x_i	u_i
0	37	
1	$19 * 37 + 11$ $mod\ 833 = 714$	$\frac{714}{832} = 0.858$
2	$19 * 714 + 11$ $mod\ 833 = 249$	$\frac{249}{832} = 0.299$
3	$19 * 249 + 11$ $mod\ 833 = 577$	$\frac{577}{832} = 0.694$

Discrete distribution

Interval method

	Probability	Return value u_i
1	0.4	1
2	0.3	4
3	0.15	4
4	0.1	8
5	0.05	16

```
x = U(0,1)
c_i = 0
For i = 1 : n
   if c_{i-1} ≤ x ≤ c_i
      return value = u_i
      break;
   end
end
```

Continuous distribution

Inversion sampling
1. Create $x = U[0,1]$
2. Calculate $y = F^{-1}(x)$

$$F(x) = y = 1 - e^{-\lambda x}$$
$$1 - y = e^{-\lambda x}$$
$$\ln(1-y) = -\lambda x$$
$$x = g(y) = -\frac{\ln(1-x)}{\lambda}$$

Rejection sampling

???

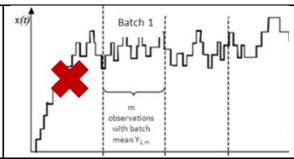
Results analysis

Simulation results are also random numbers! ->It is not enough to represent simulation results via a simple average value. We need **confidence intervals** in order to indiate the precision of our statements.

Confidence intervals are computed for desired user-defined confidence levels; for instance 95%. A confidence interval computed at 95% confidence level means that if the simulation experiment is repeated an infinite number of times, 95% of the times the estimated parameter will be inside the confidence interval

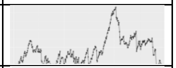
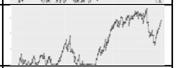
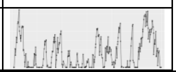
Warm-up period

Most of the times, we are interested in the long-run system behavior. The simulation run however starts with some system state which might not be realistic at all (think e.g. idle queuing systems). Consequently, we must remove this warm-up phase from the data before starting our analysis.

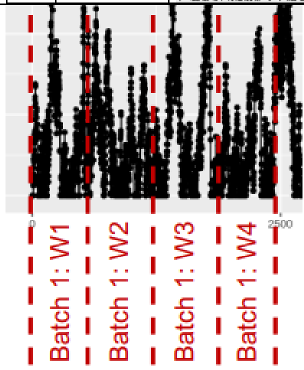


Methods for calculating Confidence Interval (CI) in DES

Independent simulation runs
In this method, multiple simulation runs are carried out, each with its own random seed. Each simulation run yields one sample mean Xi. By design, all Xi are uncorrelated. This means that we can apply basis statistics. The disadvantage of this method is that in each experiment we need to throw away the first 1 – 2 batches. Also form an administrative point of view, this approach is not ideal.

	Seed	
1	233245	
2	546815	
...		
X	321584	

Batch-Means method
After the removal of the warm-up period, simulation time is divided into n equal intervals. For each interval the mean value for the desired performance measure is computed. Under the assumption that the batches are independent (which is the case if each batch is long enough), the confidence interval for the estimated parameter can be computed from the empirical variance of the batch means.



$$\left[\hat{\theta} - z_{\frac{\alpha}{2}}\sqrt{\frac{S^2}{n}};\hat{\theta} + z_{\frac{\alpha}{2}}\sqrt{\frac{S^2}{n}}\right]$$

$\hat{\theta}$	Sample mean
$z_{\frac{\alpha}{2}}$	Constant
S	Sample variance
n	# of Batches

α	$z_{\alpha/2}$
0.1	1.645
0.05	1.960
0.01	2.576

Valid if $n \geq 30$

Formula for sample mean

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_i$$

Formula for sample variance

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\hat{\theta}_i - \hat{\theta}\right)^2$$

$\hat{\theta}_i$: mean of batch i

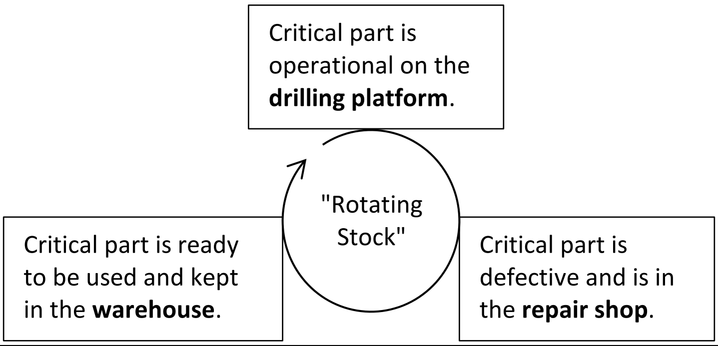
Repairable systems

Consider highly complex systems such as aircrafts, oil drilling and nuclear power plants. These systems contain thousands of critical, expensive components. To avoid enormous costs that result from system failures, such systems keep an inventory of spare parts which are flown directly to the interruption site (for instance to the oil drilling platforms). The defective parts can be repaired at an on-site repair shop.

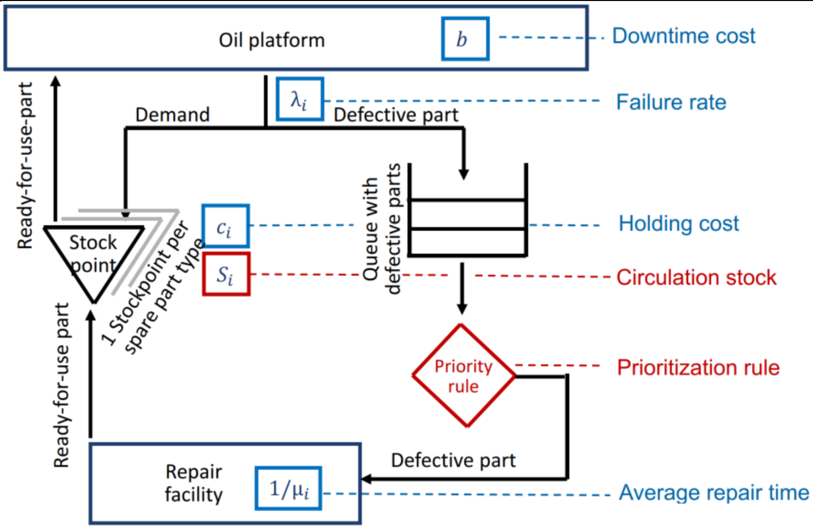
Assumptions

- Each component failure makes the system go down
- The technical systems and its spare parts are produced in a single batch. Reordering is not possible/allowed.
- Failed components can always be repaired and have the same lifetime distribution as new parts.
- The repair center can be modeled as a single server queue, only one item can be repaired at the same time.
- Transportation times between oil platfom, stocking location, and repair center can be ignored.

Workflow



Flow-Diagram

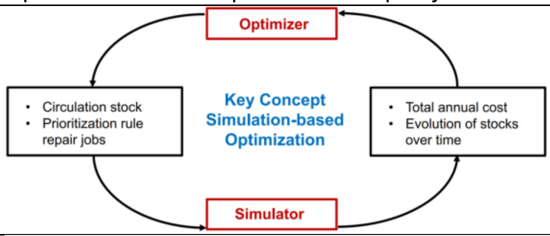


Objective
Minimize the sum of annual downtime costs, annual amortized purchasing costs and annual inventory holding costs.

Decision

Strategic: How many items do we purchase (produce) at time $t = 0$?
Operational: How to priorize the repair jobs at the repair center?

Simulation-based optimization



Optimizer finds promising stock level vectors, Simulator evaluates.

Process logic

time	Event	In Repair	Stock level Part A	Stock level Part B	Failed items Part A	Failed items Part B	# Oil platforms up	# Oil platforms down
0			2	1	0	0	100	0
4	Failure A_1	$\varnothing \rightarrow A_1$	1	1	1	0	100	0
6	Failure A_2	A_1	0	1	2	0	100	0
7	Failure B_1	A_1	0	0	2	1	100	0
8	Repair ready	$A_1 \rightarrow A_2$	1	0	1	1	100	0
10	Failure B_2	A_2	1	0	1	2	99	1 (1xB)
12	Repair ready	$A_2 \rightarrow B_1$	2	0	0	2	99	1 (1xB)
15	Failure A_3	B_1	1	0	1	2	99	1 (1xB)
16	Repair ready	$B_1 \rightarrow B_2$	1	0	1	1	100	0
20	Repair ready	$B_2 \rightarrow A_3$	1	1	1	0	100	0
24	Repair ready	$\varnothing$	2	1	0	0	100	0

start stock of [2 1]
repair time of 4 time units

"In repair" shows which spare part is currently bein repaired.

"# oil platforms" also holds the spare part that is causing the system downtime.

$$\#platform\ downs = \max(failed - stock, 0)$$

Probability $f_i(k)$

k	$f_A(k)$	
0	7/24	0,1,2,3, 12,13,14
1	15/24	4,5, 8,9,10,11, 15-23
2	2/24	6,7
3	0	-

k	$f_B(k)$
0	11/24
1	7/24
2	6/24
3	0

simulation horizon is too small to make reliable estimates on $f_i(k)$.

Optimization under FCFS	Increase for each spare part the circulation stock as long as the marginal savings $[1 - F(S)]\, b$ exceed the marginal (investment) c_i. $$S_i^* = \min\left\{Q \middle\| F_i(Q) \geq \frac{b - c_i}{b}\right\}$$	b: downtime cost $40 * 365d$ c: amortisation/rent 730 $\frac{40 * 365 - 730}{40 * 365} = 0.95$
Dynamic Prioritization	1) Spare parts causing downtime cost at this very moment, are prioritized. 2) If multiple spare are causing downtime cost, the one with the smallest average repair time will be selected (to continue as soon as possible) 3) When all oil platforms are running, we pick the spare part with the lowest "coverage". $= (current\ stock + 1) * average\ time\ between\ 2\ failures$ **Heuristcs** 1. Circulation stock vector S influences repair priorities. 2. Repair priorities determine sequence of repair activities. 3. Sequence of repair priorities influence $f_i(k)$, the distribution of defective parts of spare part i over time. 4. The distributions $f_i(k)$ are crucial inputs for the calculation of circulation stocks according to $S_i^* = \min\left\{Q \middle\| F_i(Q) \geq \frac{b-c_i}{b}\right\}$	
	Iterative method for calculating S^* **Step 1a**: Develop (or use) an intelligent Prioritization rule Ω **Step 1b**: Initialize iteration count n=0 und circulation stock vector $\boldsymbol{S^n} = \boldsymbol{0}$ **Step 2a**: Simulate the system under Ω und $\boldsymbol{S^n}$. Save the relative frequencies $i,\ f_i(k), k \geq 0$. **Step 2b**: Calculate $\boldsymbol{S^{n+1}} = \min\{\,Q \mid F_i(Q) \geq \frac{b-c_i}{b}\}$ **Step 2c**: Stop when $\boldsymbol{S^{n+1}} = \boldsymbol{S^n}$. Otherwise n = n + 1 and go to Step 2a.	