Overview of the ABM Process & ODD Protocol

Nathaniel Osgood
MIT 15.879

February 29, 2012

Model Specification Mechanisms

Stock & Flow Models: "Hedgehog Knowledge"

- Small modeling vocabulary
- Power lies in combination of a few elements
- Analysis conducted predominantly in terms of elements of model vocabulary

Agent-Based Modeling: "Fox Knowledge"

- Large modeling vocabulary
- Different subsets of vocabulary used for different models
- Power in flexibility & combination of elements
- Variety in analysis focus

ABMs: Larger Model Vocabulary & Needs

- Events
- Multiple mechanisms for describing dynamics
  - State diagrams
  - Stock and flow
  - Custom update code
- Inter-Agent communication (sending & receiving)
- Multiple types of transitions
- Diverse types of agents
- Spatial & topological connectivity & patterning
- Subtyping
- Mobility & movement
- Graphical interfaces
- Data output mechanisms
- Stochastics complicated
  - Scenario result interpretation
  - Calibration
  - Sensitivity analysis
- Synchronous & asynchronous distinction, concurrency

The Overview, Design concepts, and Details (ODD) Protocol for ABM Design

- Consensus protocol derived from panel fo ABM modelers
- Primary focus: *Specification* protocol
  - To help understand, communicate & reproduce ABMs
- Secondary benefit: Process for ABM design

ODD: 3 Broad Components

- Overview: model goals & high level scope & design
- Design concepts: Different aspects of design being considered
- Remaining elements

Overview of Modeling Process

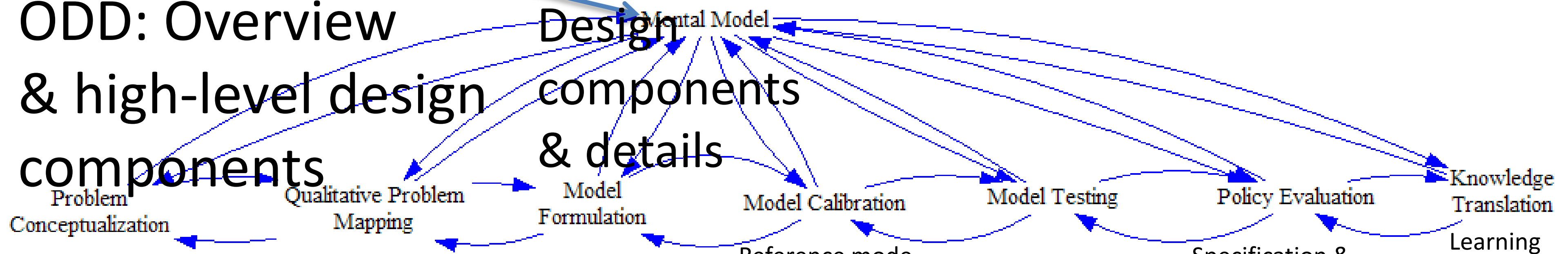
- Typically conducted with an interdisciplinary team
- An ongoing process of refinement
- Best: Iteration with modeling, intervention implementation, data collection
- Observation:
  - Traditionally, the focus in ABM has been on insights gained from the model delivered
  - Often it is the *modeling process* itself – rather than the models created – that offers the greatest value

ABM Modeling Process Overview

A Key Deliverable!

ODD: Overview & high-level design components

ODD: Design components & details



Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogeous, exogenous,ignored)

Key entities & their relationships
- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

State charts
Parameter & state variables
Qualitative State charts
Influence & Causal loop diagrams
Multi-agent interaction diagrams
Process flow structure
Key events
Rough model time & spatial extent
Observers to measure outputs

Specification of
- Parameters
- Quantitative causal relations
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds

Structural sensitivity analysis

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered
  - Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

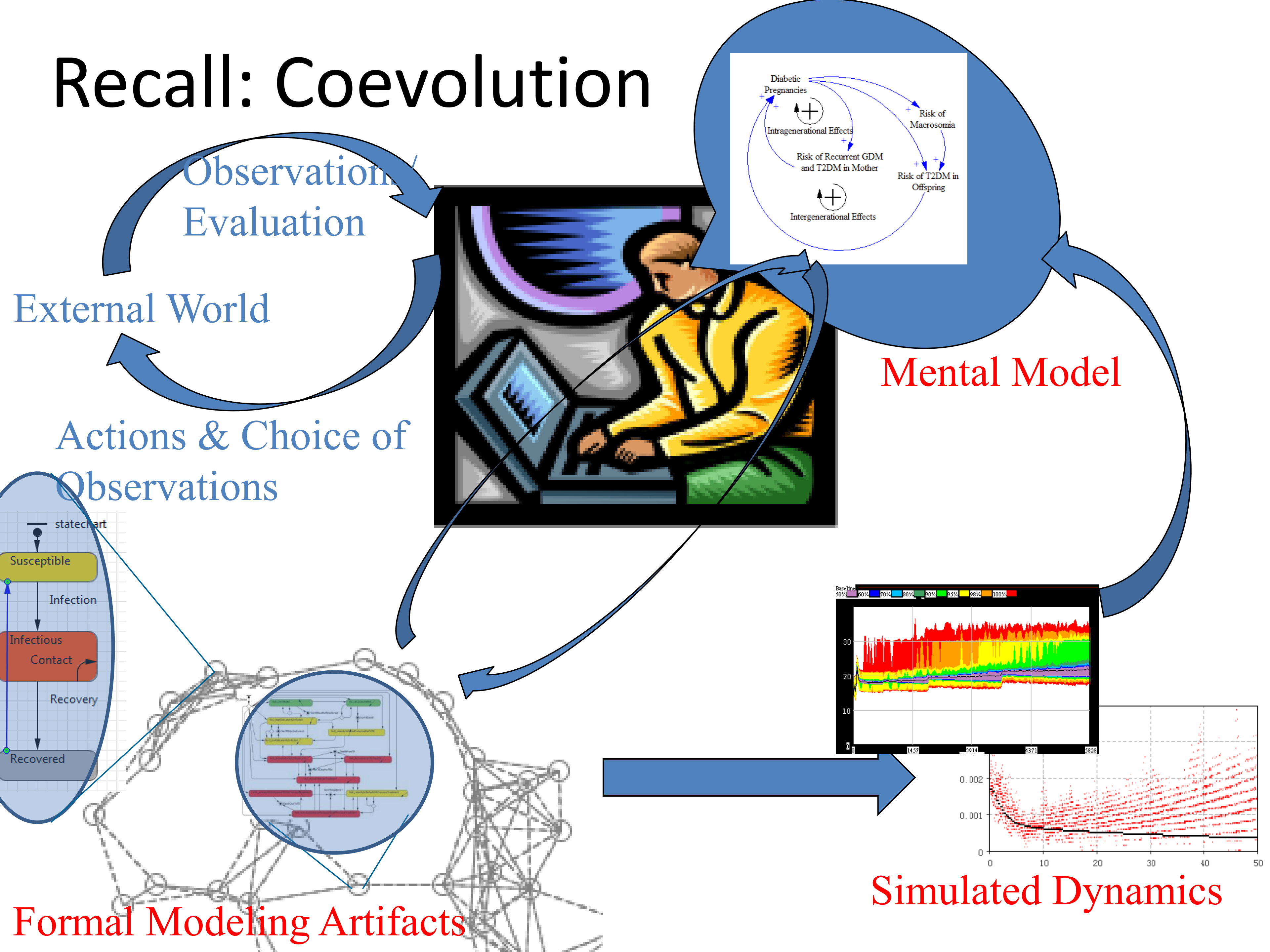
Specification & investigation of intervention scenarios
Investigation of hypothetical external conditions

Cross-scenario comparisons (e.g. CEA)

Learning environments (e.g. DISimS)

Visualizations

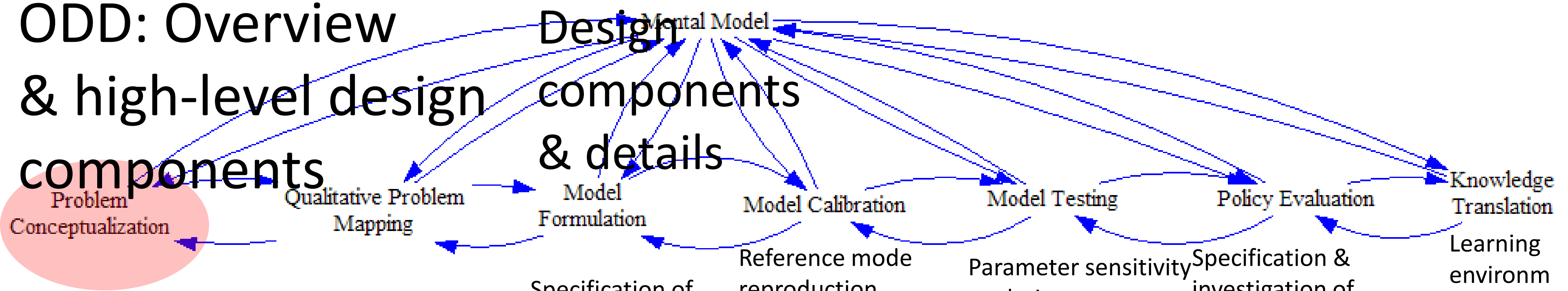
Recall: Coevolution



ABM Modeling Process Overview

ODD: Overview & high-level design components

ODD: Design components & details



Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogeous, exogenous,ignored)

Key entities & their relationships
- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

State charts
Parameter & state variables
Qualitative State charts
Influence & Causal loop diagrams
Multi-agent interaction diagrams
Process flow structure
Key events
Rough model time & spatial extent
Observers to measure outputs

Specification of
- Parameters
- Quantitative causal relations
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds
Structural sensitivity analysis

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered
  - Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

Specification & investigation of intervention scenarios
Investigation of hypothetical external conditions

Cross-scenario comparisons (e.g. CEA)

Learning environments (e.g. DISimS)

Visualizations

ODD Overview: model goals & high level scope & design

- Purpose
- Definition of key elements during operation
  - Entities
  - States
  - Scales
- Process overview and scheduling

ODD Overview: model goals & high level scope & design

- Purpose
- Definition of key elements during operation
  - Entities
  - States
  - Scales
- Process overview and scheduling

Identification of Questions/ “The Problem”

- All models are simplifications and “wrong”
- Some models are useful
- Attempts at perfect representation of “real-world” system generally offer little value
- Establishing a clear model purpose is critical for defining what is included in a model
  - Explaining reference modes
  - Understanding broad trends/insight?
  - Understanding policy impacts?
  - Ruling out certain hypotheses?
- Think explicitly about model boundaries
- Adding factors often does not yield greater insight
  - Often simplest models give greatest insight
  - Opportunity costs: More complex model takes more time to build=>less time for insight

Importance of Purpose

Firmness of purpose is one of the most necessary sinews of character, and one of the best instruments of success. Without it genius wastes its efforts in a maze of inconsistencies.

Lord Chesterfield

The secret of success is constancy of purpose.

Benjamin Disraeli

The art of model building is knowing what to cut out, and the purpose of the model acts as the logical knife. It provides the criterion about what will be cut, so that only the essential features necessary to fulfill the purpose are left.

John Sterman

[W]e keep [ABM] models simple by using the question or problem addressed with the model as a filter. Aspects of the real system…should be included in the model only if we, or the experts we are working with, consider them absolutely essential for solving a particular problem

Railsback & Grimm

Adapted from H Taylor, 2001

Model Purpose & Agent-Based Models

- The flexibility & generality & computational universality of ABM supports the creation of arbitrarily rich models
- Typically: $\exists$ high opportunity cost to investing in a given model area: Given limited time, it takes away from richness elsewhere – and often from learning
- Given this flexibility & cost, it is especially critical to wield the "logical knife" of model purpose
- YAGNI (**Y**ou **A**in't **G**onna **N**eed **I**t): Start simple & add as one develops confidence in & understanding of model

Common Division

- Endogenous
  - Things whose dynamics are calculated as part of the model
- Exogenous
  - Things that are included in model consideration, but are specified externally
    - Time series
    - Constants
- Ignored/Excluded
  - Things outside the boundary of the model

Pattern Oriented Modeling

- ABMs occupy an arbitrarily rich model space
- To gain insight, it important to leverage the broad set of information we know about a system
  - We need to accompany general model purpose by a broad set of *patterns to be explained*
- Patterns (often called "stylized facts) may include e.g.
  - Similar to classic System Dynamics reference modes
    - Quantitative time series patterns
    - Qualitative (e.g. presence of oscillations, rising, asymmetries, etc.)
  - Patterns of heterogeneity (disparities, stratification, deg. dist)
  - Spatial/topological patterns (waves, clustering, phenomenology)
  - Multi-scale phenomena
- ***We seek a model that will explain (or at least exhibit consistency with, "stay true to") such patterns***

The Value of Patterns in Building a Model

- Patterns are pieces of information which – if the model didn't match them – it would cast suspicion on the model
- Typically specific to the purpose (if goal were different, we'd use a different set of patterns)
- Try to rule out possible submodels using patterns
- Try to use very broad set of knowledge
  - Even if a given pattern is "weak" in constraining the model (e.g. Ng rates higher among women than men), a set of such weak patterns can collectively greatly constrain possible dynamic hypotheses (ABM structure)

Patterns in Confidence Building

- When inspecting model results, we will seek to recognize other patterns (not built into model or used to judge it) & use them for cross-validation
- Example patterns from a model
  - Prior spatial distribution emerging from movement patterns
  - Contact patterns emerging from individual movement
  - Case-contact network structure emerging from contact tracing process

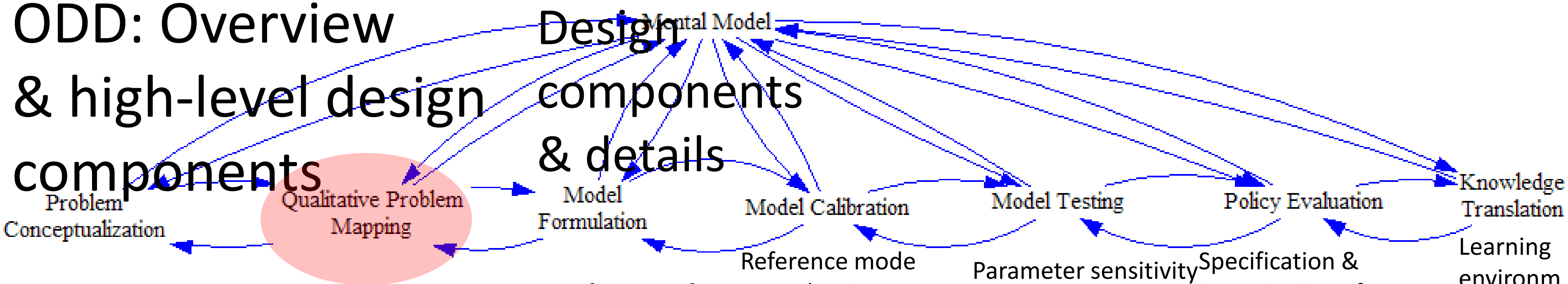
Example Phenomenological Patterns

- Flocking
- Oscillation
- Gradients
- Waves
- Cascaded transitions over time
- Phase change phenomena
- Clustering
- “Waves” of topological spread/precolation
- Punctuated equilibria

ABM Modeling Process Overview

ODD: Overview & high-level design components

ODD: Design components & details



Problem Conceptualization

Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogeous, exogenous,ignored)

Key entities & their relationships
- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

Qualitative Problem Mapping

State charts
Parameter & state variables
Qualitative State charts
Influence & Causal loop diagrams
Multi-agent interaction diagrams
Process flow structure
Key events
Rough model time & spatial extent
Observers to measure outputs

Model Formulation

Specification of
- Parameters
- Quantitative causal relations
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Model Calibration

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds

Structural sensitivity analysis

Model Testing

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered
  - Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

Policy Evaluation

Specification & investigation of intervention scenarios
Investigation of hypothetical external conditions

Cross-scenario comparisons (e.g. CEA)

Knowledge Translation

Learning environments (e.g. DISimS)

Visualizations

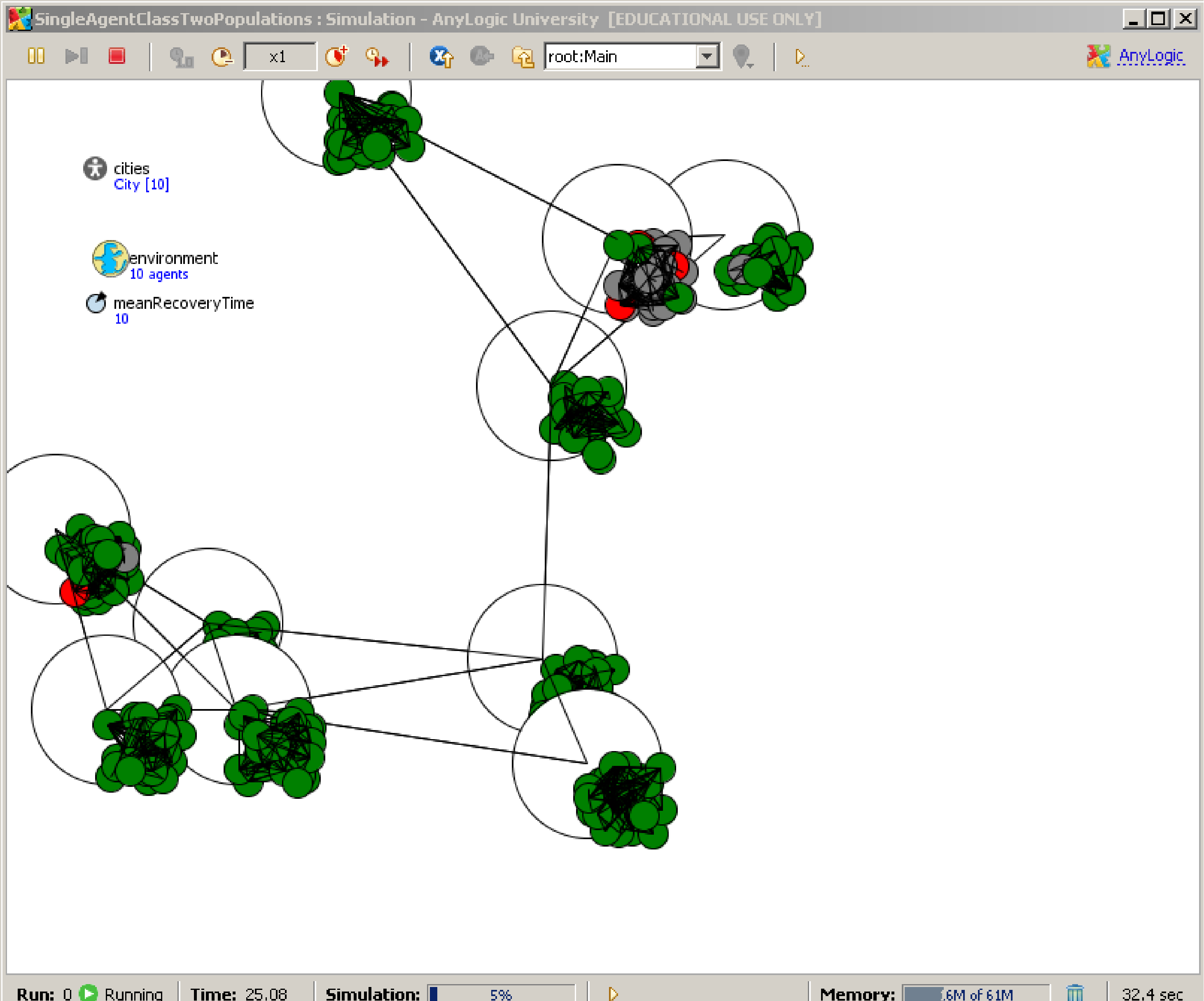
Overview: model goals & high level scope & design

- Purpose
- Definition of key operational elements
  - Entities
  - States
  - Parameters
  - Scales
- Process overview and scheduling

Definition of Key Operational Elements

- Entities & their Relationships
  - 1 or more types of agents
  - Local environment Contexts (Spatial patches, hierarchical elements) in which particular agents circulate (neighborhood, school, cities)
  - Global environment: Contexts & Processes operating across network or spatial extent of model (Seasonality, Regulations, tax code)
  - Who interacts with whom? In what capacity?
- States
- Scales (temporal & spatial)

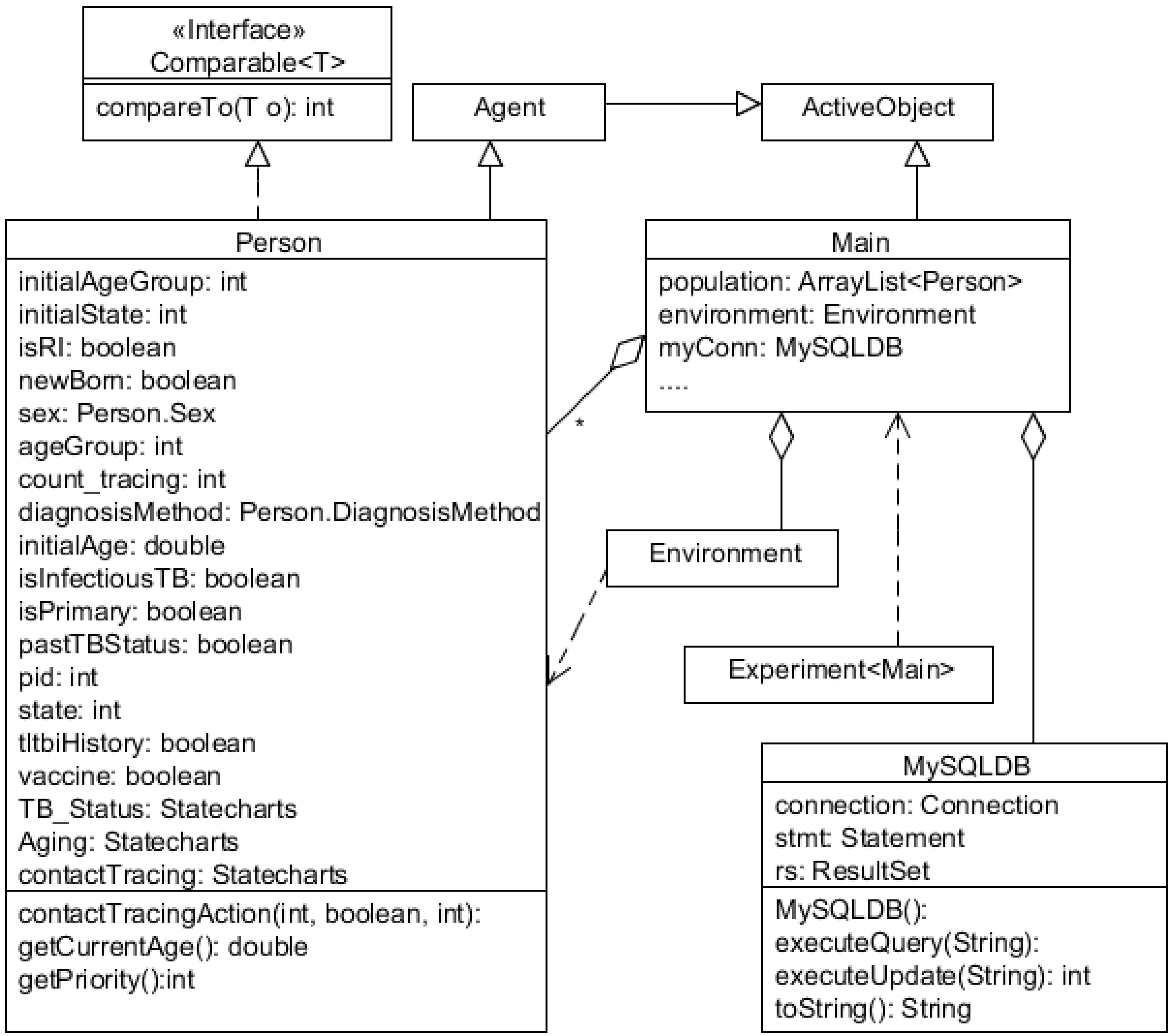
Stylized Example Relationships: Nesting, Network Linkages, Spatial Proximity



Properties (ODD: “State”)

- Parameters: Mostly fixed characteristics
  - Person: Sex, Ethnicity, Birth date & place
  - Specifying source of information used for parameter values is key for model documentation
- Evolving components: State
  - Person: Continuous: Body Weight & composition, immune status, Income, Savings[, Age] Discrete: Broad Smoking status ({Never, Current, Former})
  - Characterizes “the current situation of the system”
  - Elements of natural history of infection
- Note that this distinction is useful for conceptual clarity and in AnyLogic, but is not recognized by ODD

Documenting Agent Characteristics in UML



Subtleties

- Not part of [Represented] State: Derivative Information e.g.
  - Weight (if derived from body compartments mass)
  - Mean city income (if derived from incomes of resident population)
  - Agent age (if derived from current age and birthdate)
- Often we can elect one of many “coordinate systems” in which to describe properties e.g.
  - Weight and fractional body composition vs. body mass in different compartments
  - BMI & Height (Weight derived) vs. Weight & Height (BMI derived)

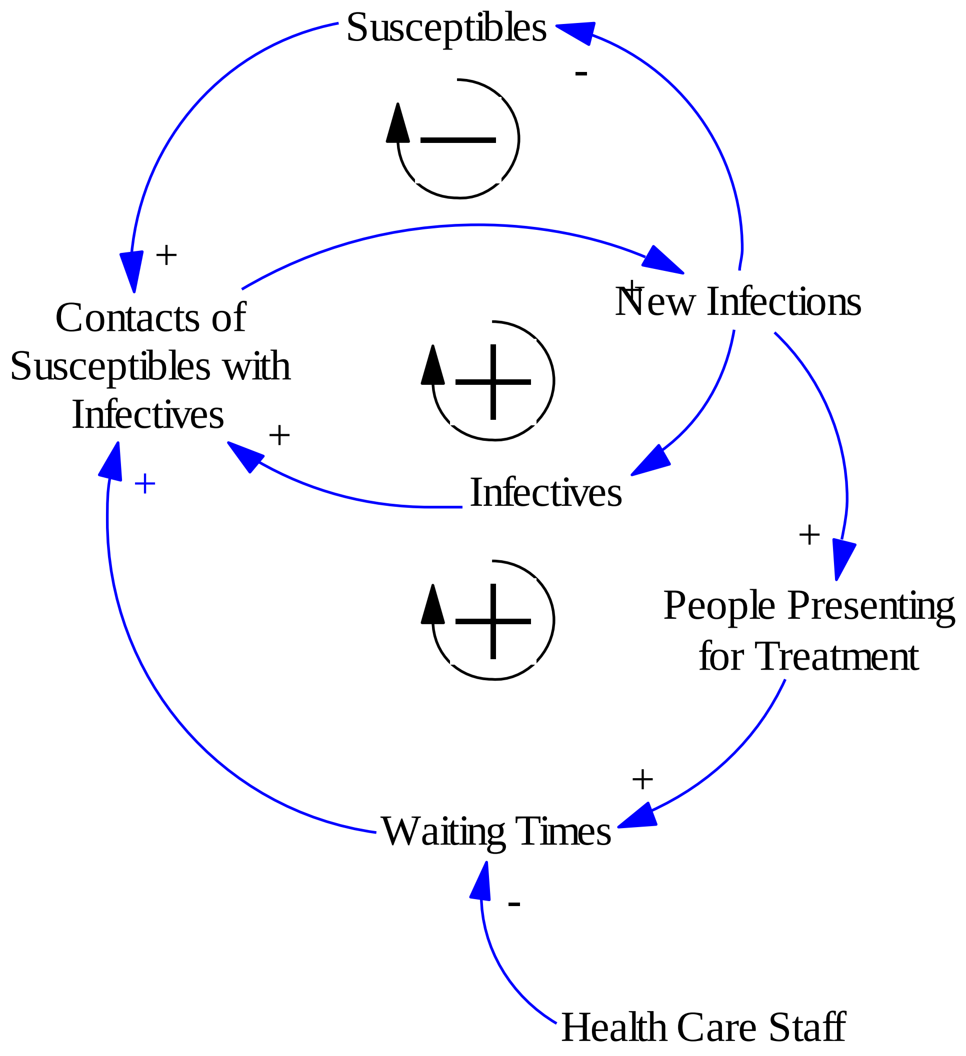
Scales

- Spatial
  - Spatial extent
  - Any spatial step size
- Time
  - Time extent
  - Any time step size
- In both cases
  - Often things are taken as homogenous within a step
  - **Extent** is typically determined by desire to capture richness of emergent phenomenon being simulated
  - **Step size** is typically dictated by entity consideration (e.g. the need to capture inter-agent interaction, agent evolution)

Local Environment

- Examples: Neighborhoods, schools, cities
- Note that each type of context may have different
  - Processes operating
  - State information (beyond subpieces)
- Neighborhood
  - Mixing
  - State: land allocation, components of built environment, food resources
- Schools
  - Processes: Admissions & graduation & gating processes, class assignment, remediation problems
  - State: Mapping from student => teacher, student => classses
- Cities
  - Processes: Municipal services
  - State: Tax records, etc.

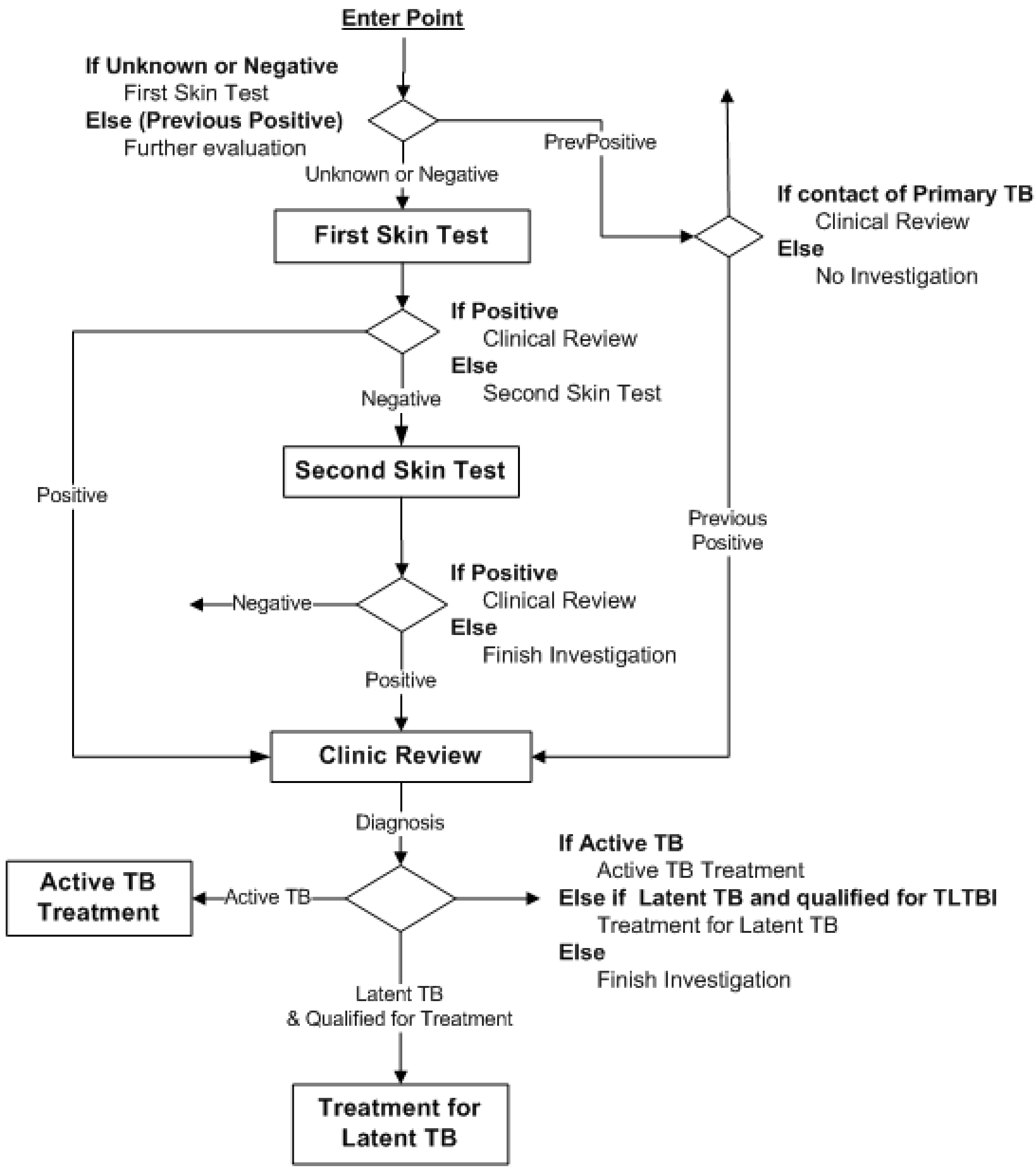
Reminder: Causal Loop Diagram



Process Overview & Scheduling

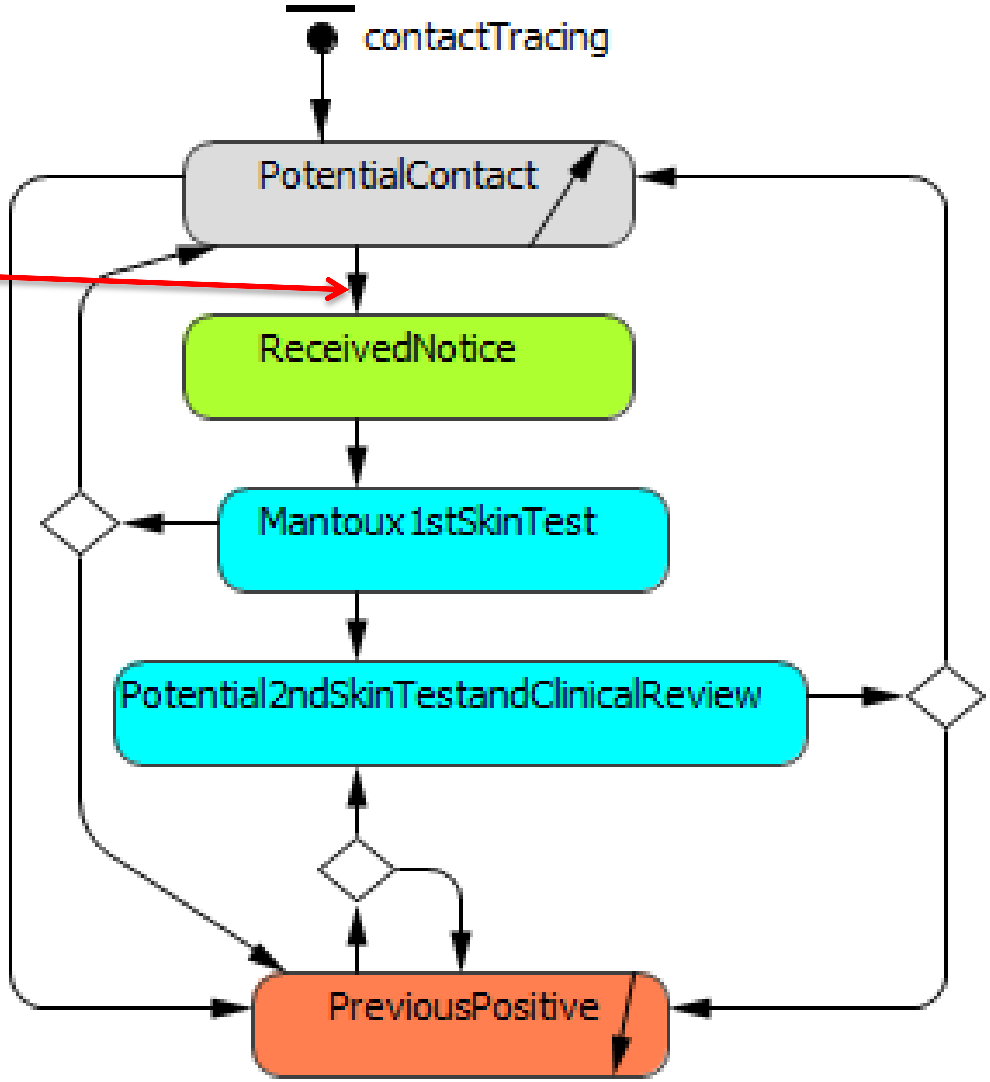
- Here, we are dealing with the dynamics of the model
- Most: Entity-related processes [1]
  - Use list of entities to guide analysis: Specify WHAT "is happening" (is "going on") for different entities
    - In isolation
    - In their interactions
  - Abstraction is key here -- can specify generic term, and leave the details for a submodel
- Observer processes
  - These collect information on the model (e.g. for reporting)

[1] NB: “process” is being used here in a mathematical sense



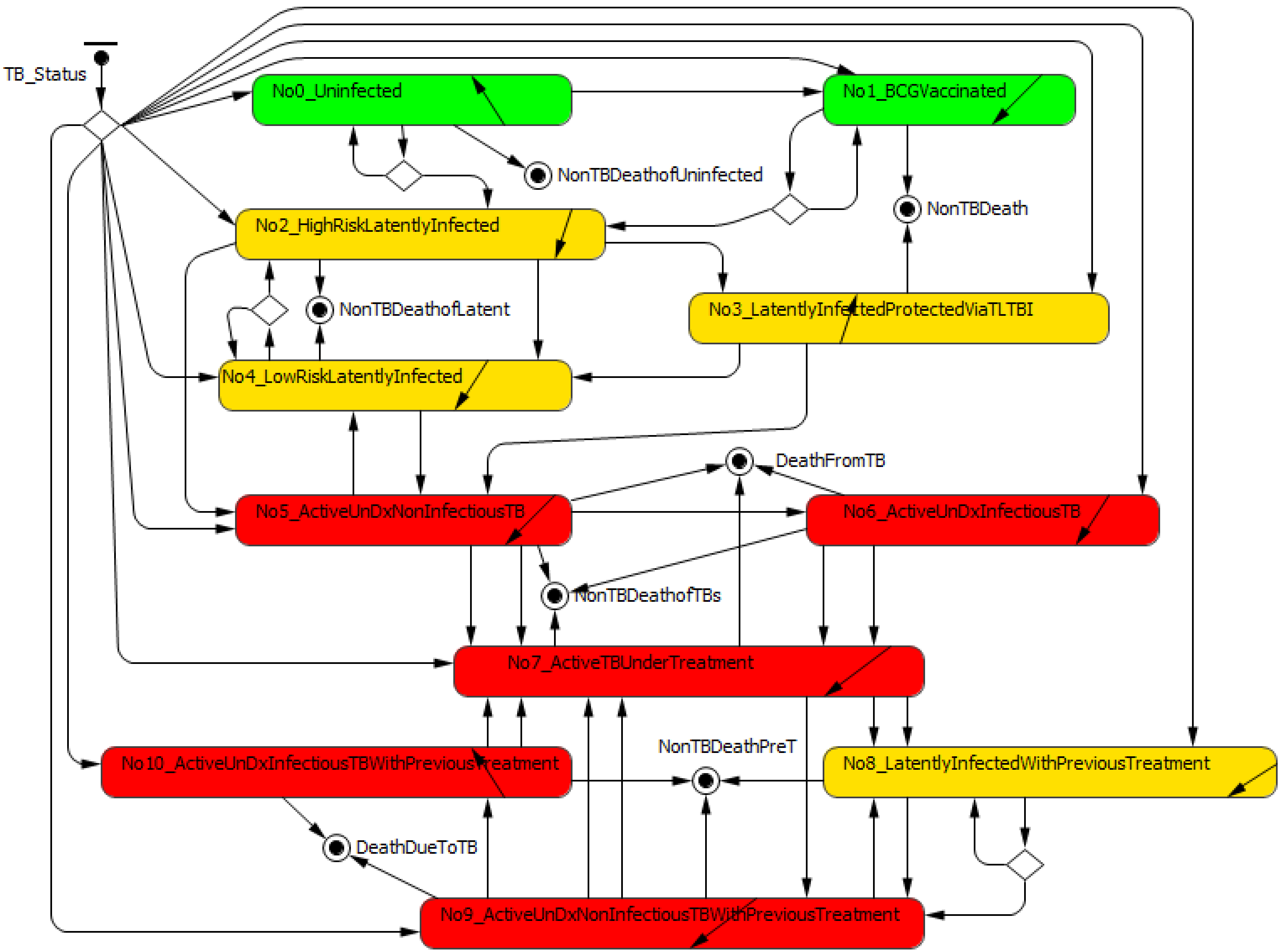
Note: Contact tracing for primary case is different. Since the target is to find the source of infection, once the presumed source is found, contact tracing is discontinued. If the contact is not positive in a TST test, then no need for the 2nd skin test.

State Diagrams to Document State Progression

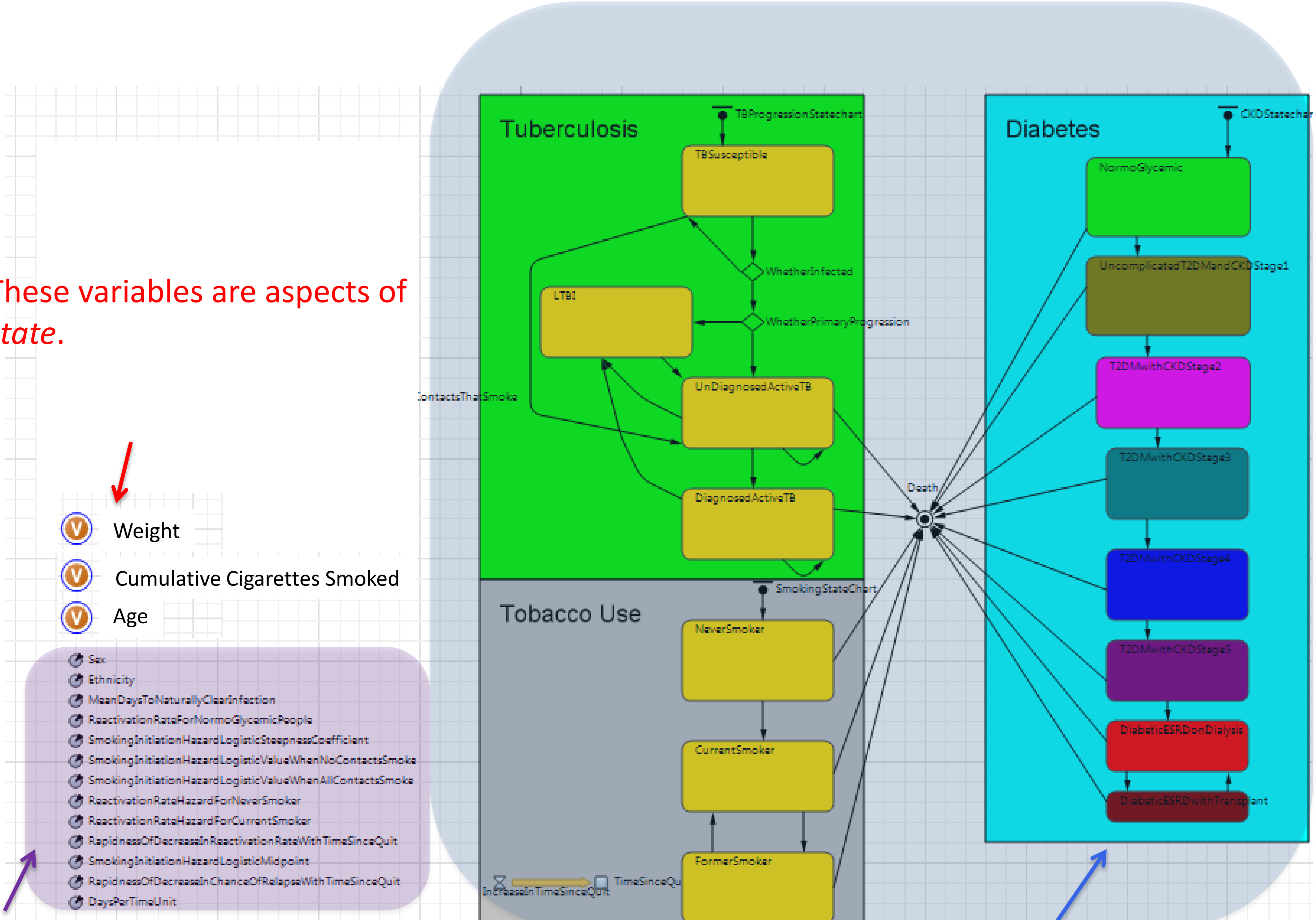


Qualitative Transitions (no likelihood or governing rules yet unambiguously specified)

State Diagram 2



These variables are aspects of *state*.



These "parameters" give static characteristics of the agent

These describe the "behaviours" – the mechanisms that will govern agent dynamics