Probabilistic Inferences in Compositional Hierarchies

Dagstuhl Workshop
"Logics and Probabilities in Scene Interpretation"
24.- 29.2.2008

Bernd Neumann
Cognitive System Laboratory
Hamburg University

Introduction

We try to understand scene interpretation and to implement a generic scene interpretation system

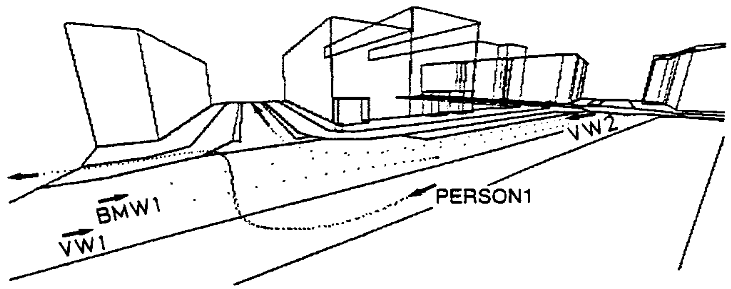
Summary

- **Scene interpretation can be formulated as stepwise "partial model construction" based on a compositional concept hierarchy**
- **Probabilities provide guidance for partial model construction**
- **Probabilistic inference can be efficiently realized for a compositional hierarchy with abstraction properties**

Animated Slide!

Background

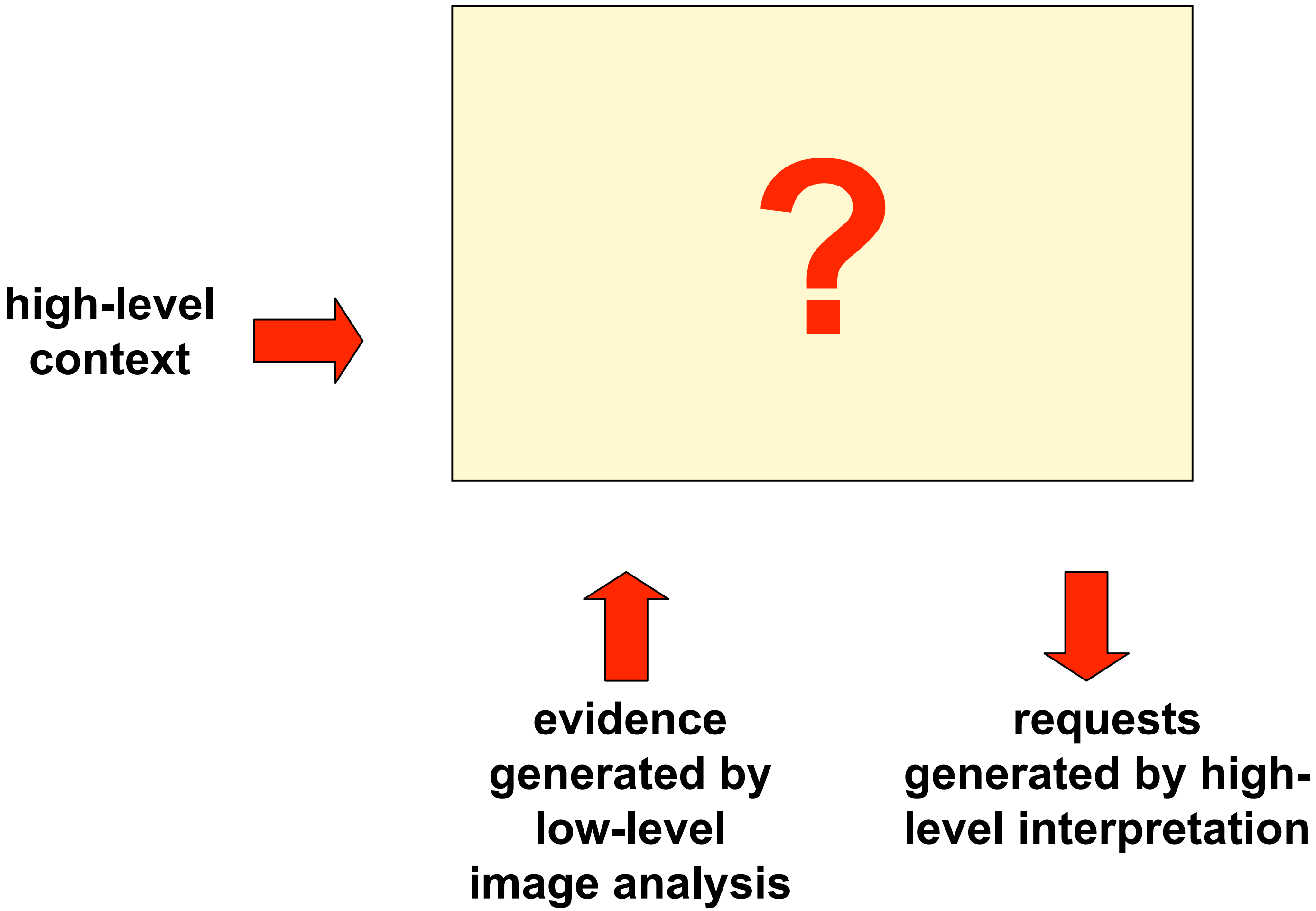
- Early work on interpretation and natural-language description of traffic scenes (1981-89)



- Interpreting table-laying scenes using configuration technology (2001-04)

- Learning man-made structures and interpreting building facades (2006-09)

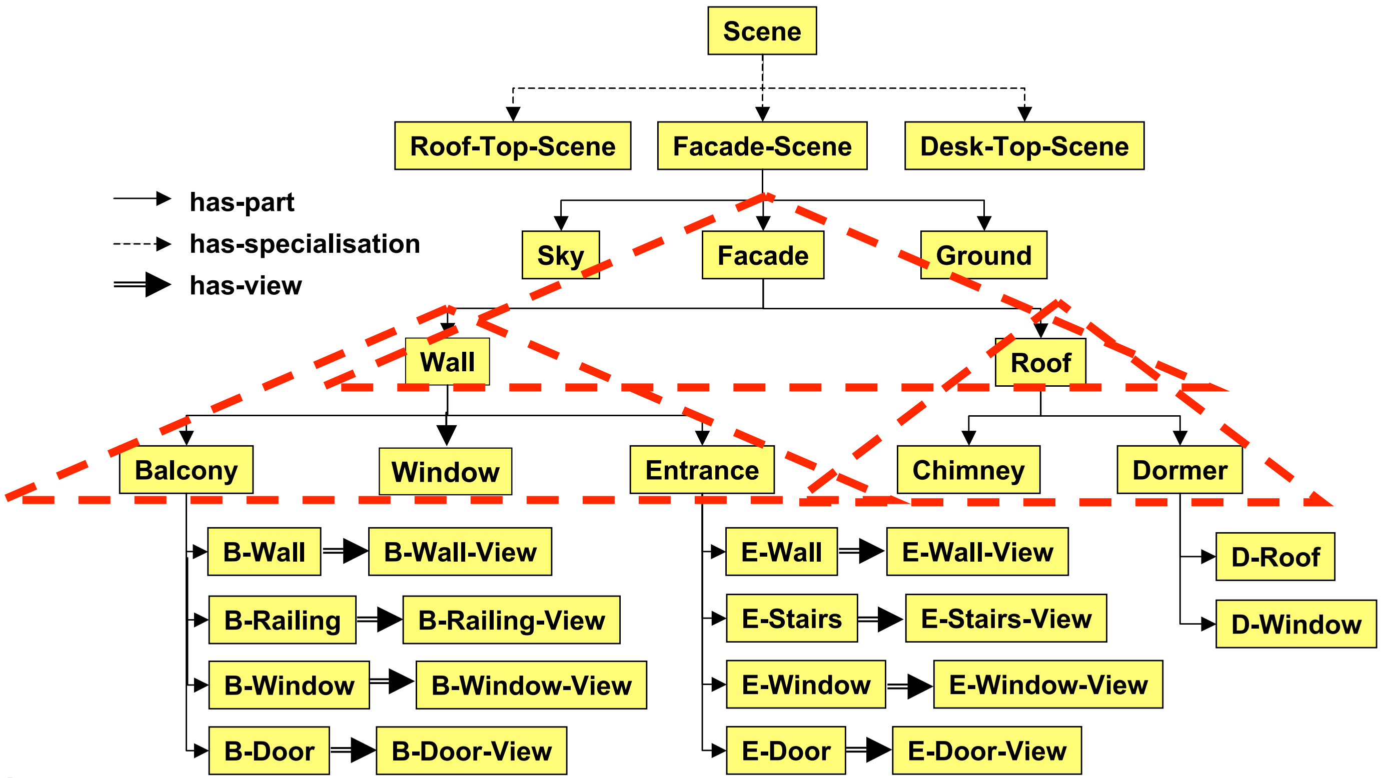
High-level Knowledge in a Scene Interpretation System



Structuring High-level Knowledge

To interface with human concepts and common knowledge, a generic approach requires:

- **Object-centered representations**
- **Compositional hierarchies with abstraction => aggregates**



Animated Slide!

Aggregate Structure in SCENIC

"SCENIC" = Scene Interpretation by Configuration

<u>Example</u>: Concept for horizontally aligned facade objects (e.g. window array)

:name	Hor-Formation	
:parent	Formation	
:parameters	Y-Difference	[0 5]
	Low-Left-X	[0 INF]
	Low-Left-Y	[0 INF]
	Up-Right-X	[0 INF]
	Up-Right-Y	[0 INF]
:relations	has-elements	{Facade-Object [2 inf]}
	element-of	{Facade [1 inf]}
	left-of	{Physical-Object [0 inf]}
	right-of	{Physical-Object [0 inf]}
	above	{Physical-Object [0 inf]}
	under	{Physical-Object [0 inf]}
:contraints	Y-Alignment-Constraint	

Y-Alignment-Constraint on Y-Difference and Facade-Object coordinates

constraints defined elsewhere

Instantiated Aggregate

:name	Hor-Formation1	
:instance-of	Hor-Formation	
:parameters	Y-Difference	[0 5]
	Low-Left-X	[0 100]
	Low-Left-Y	[50 55]
	Up-Right-X	[200 INF]
	Up-Right-Y	[80 85]
:relations	has-elements	{Facade-Object [2 inf] Window1 Window2}
	element-of	{Facade [1 inf] Facade1}
	left-of	{Physical-Object [2 inf] Door1, Formation-X2}
	right-of	{Physical-Object [0 inf]}
	above	{Physical-Object [0 inf]}
	under	{Physical-Object [0 inf]}
:contraints	Y-Alignment-Constraint	

Instantiated Y-Alignment-Constraint on Y-Difference and Facade-Object coordinates

instantiated constraints defined elsewhere

Scene Interpretation by Model Construction

An interpretation I = [D, φ, π] of a logical language maps
- constant symbols of the language into individuals of a real-world domain D
- N-ary predicate symbols of the language into predicate functions over D^N

A model of some clauses is an interpretation for which all clauses are true.

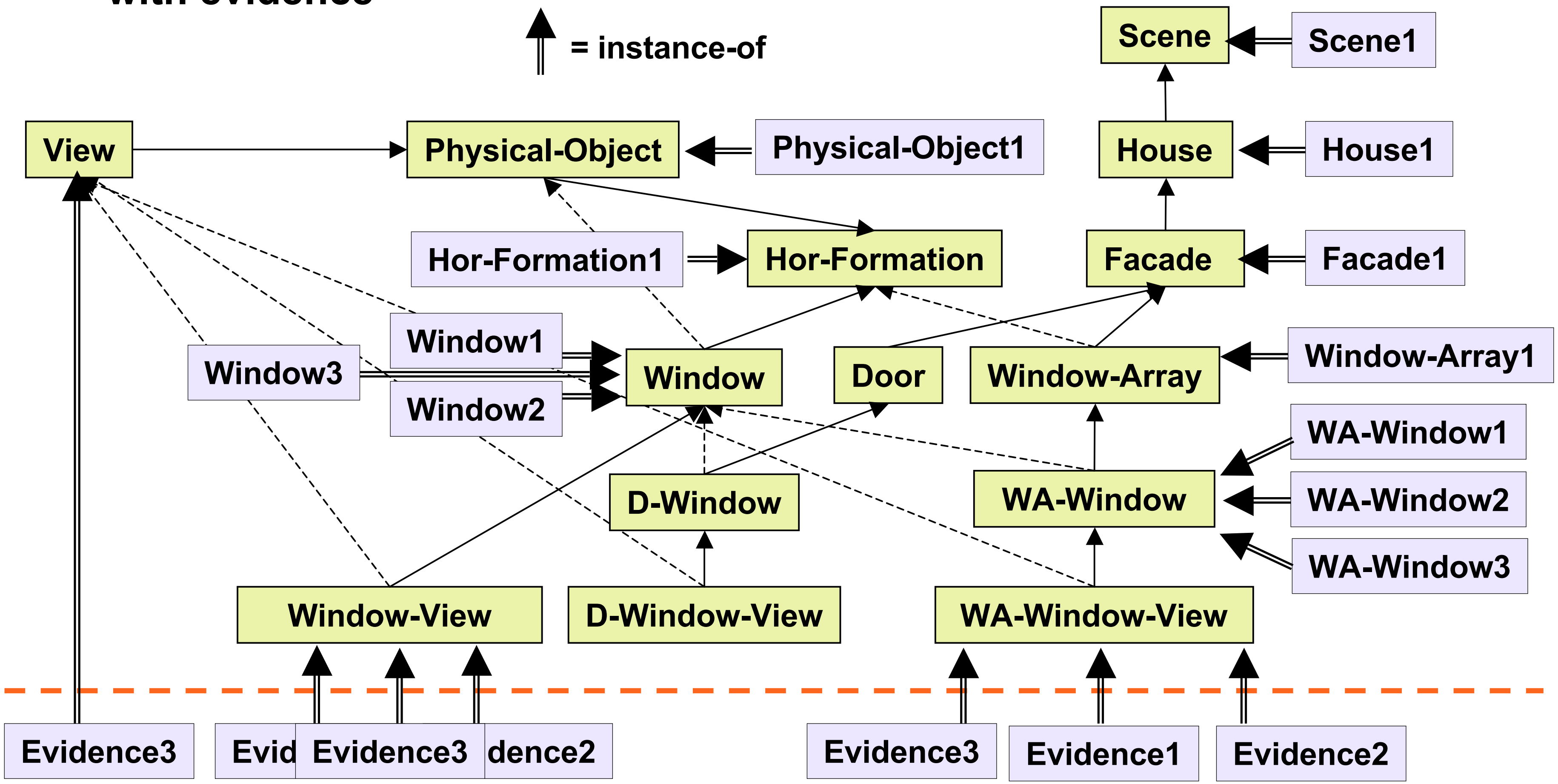
How to do model construction:

- Establish mapping φ by assigning segmentation results to constant symbols and hypothesizing other necessary constant symbols
- Establish mapping π by assigning computational procedures to predicate symbols
- Construct model by finding clauses which are true

Deciding whether a model exists is undecidable in FOPC!
There may be infinitely many models!

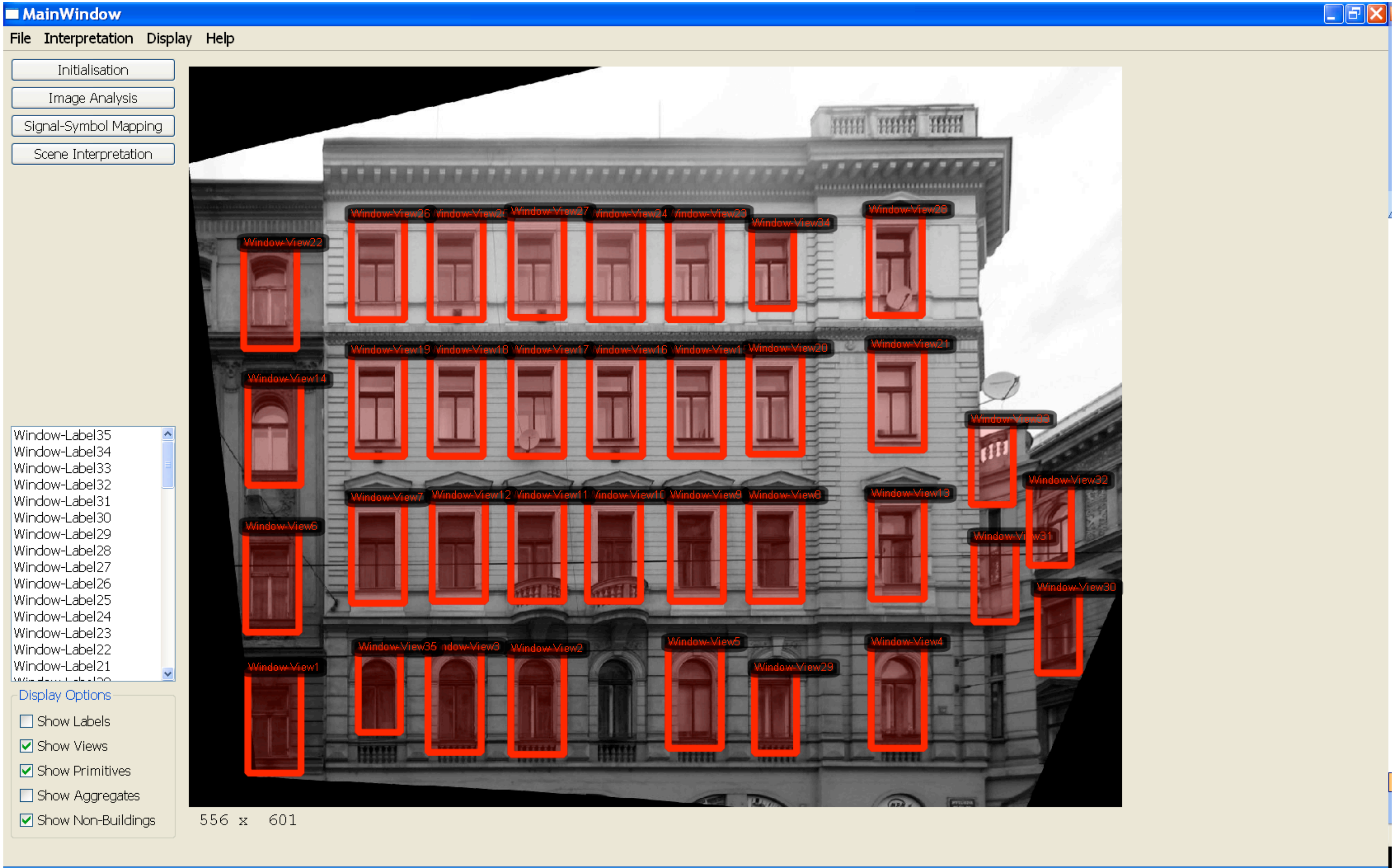
Interpretation Process

- Image analysis generates evidence
- Interpretations are stepwise instantiations of scene concepts consistent with evidence

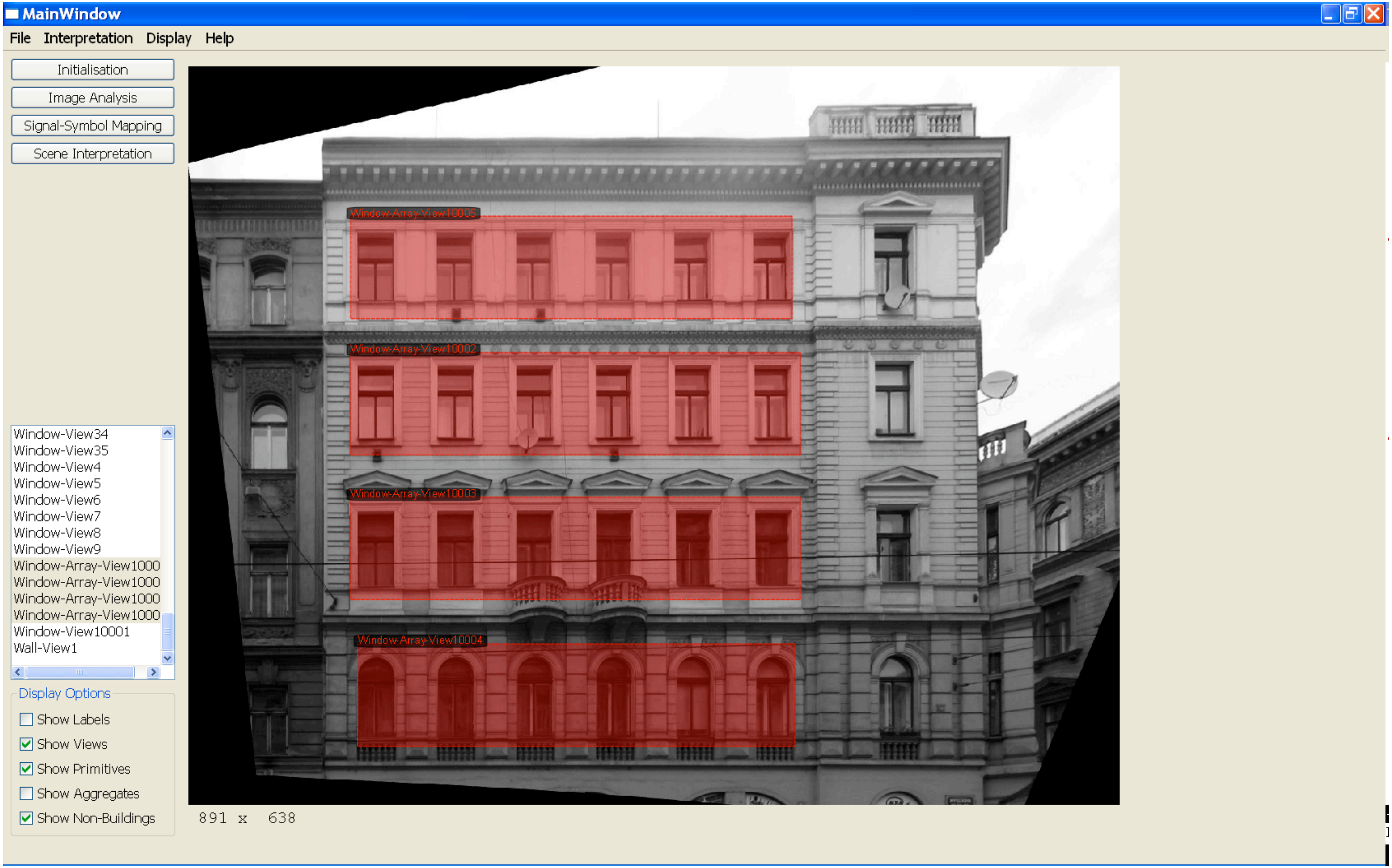


Animated Slide!

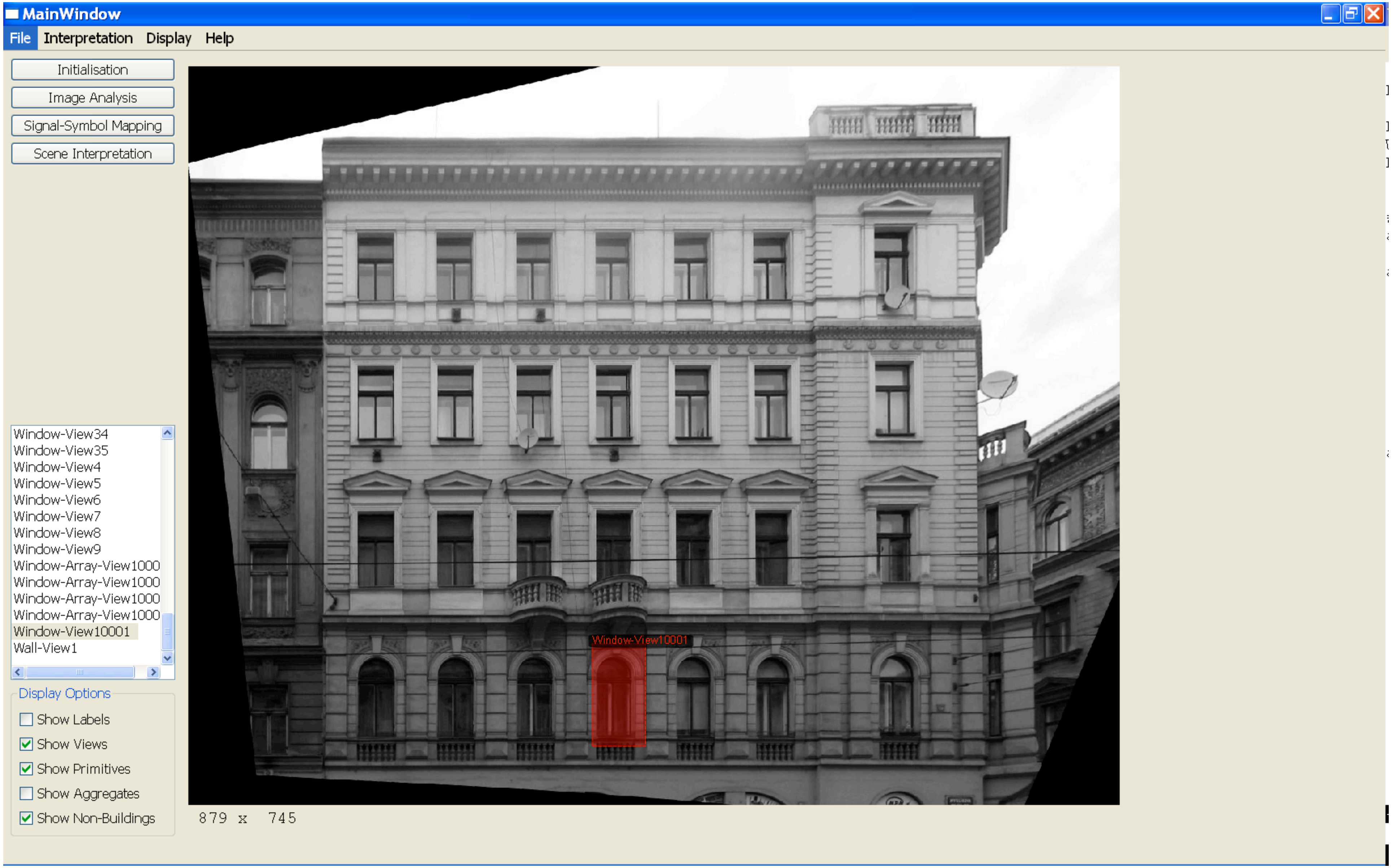
Low-level Results



Recognized Window-Arrays

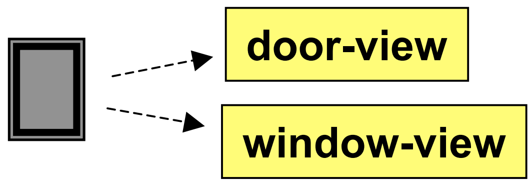


Hypothesized and Verified Additional Window

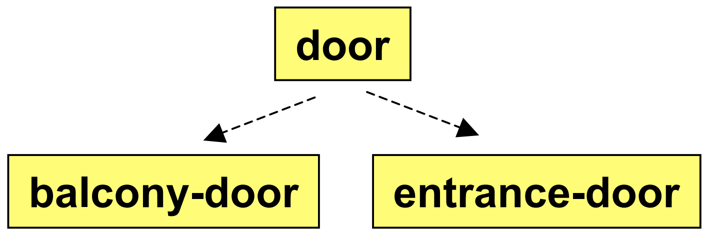


Uncertain Decisions in Stepwise Scene Interpretation

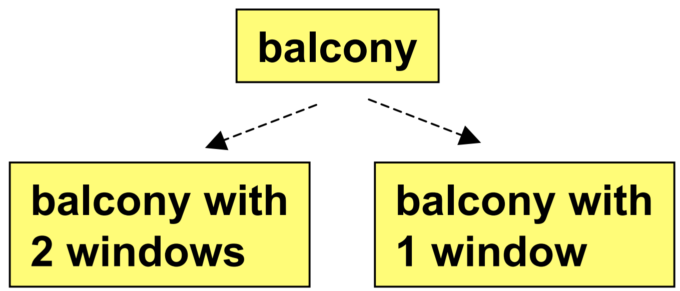
- **Evidence assignment to object views**



- **Choice of alternative specializations**

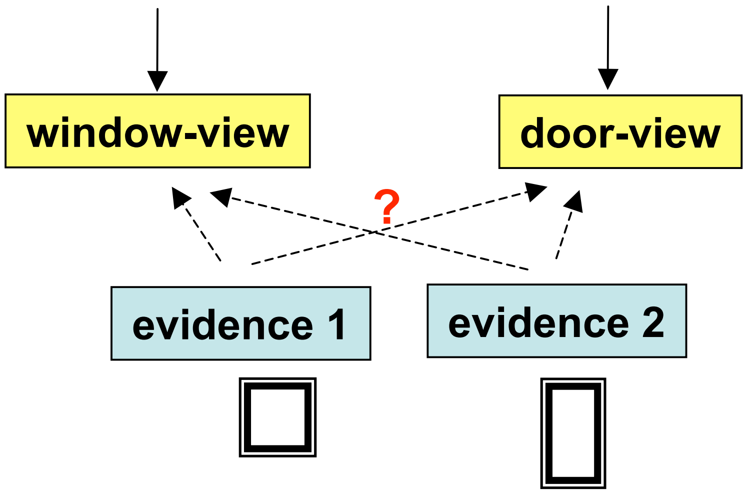


- **Choice of cardinalities in aggregate models**



Evidence Assignment Problem

To which part of an aggregate should a given evidence be assigned?



Optimal decision would require

- **postponing classification until all evidence is available**
- **maximization over all reasonable evidence permutations**

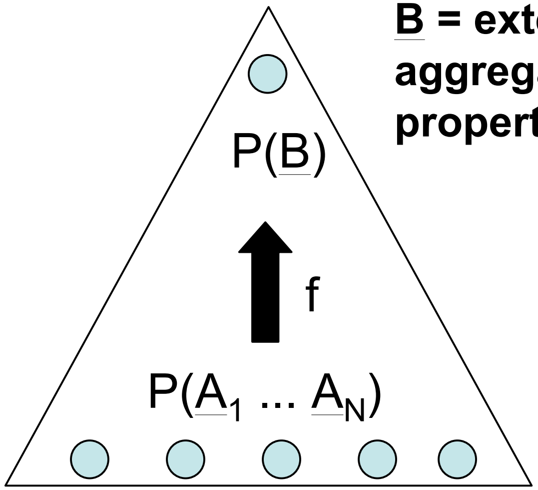
Assignment problem not encountered in Bayesian decisions or belief system reasoning!

Frequentist Probabilistic Model

Basic view:

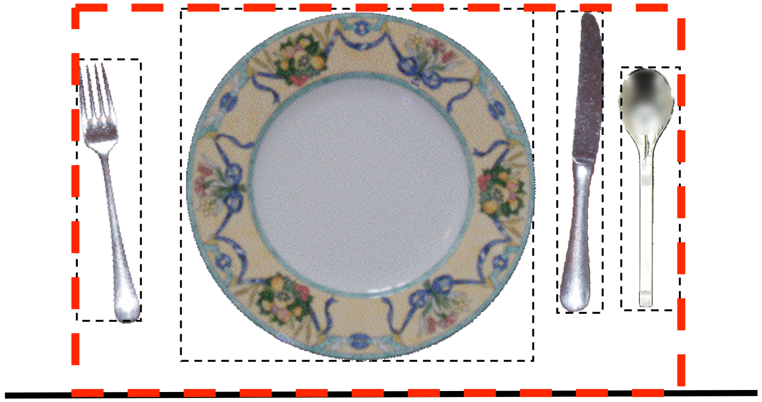
An aggregate

- **is a set of correlated parts which together constitute a meaningful entity**
- **specifies an abstraction from the descriptions of its parts**



B = external aggregate properties

Example: Bounding-box abstraction

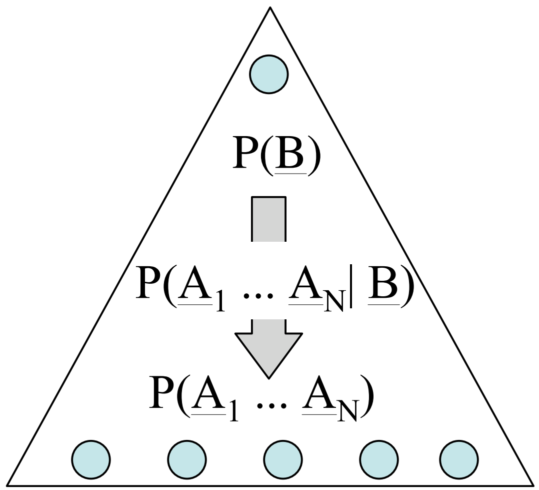


$A_1 \dots A_N$ = internal parts properties

There exists a functional mapping f : $A_1 \dots A_N$ => B

Probabilistic Aggregate Structure

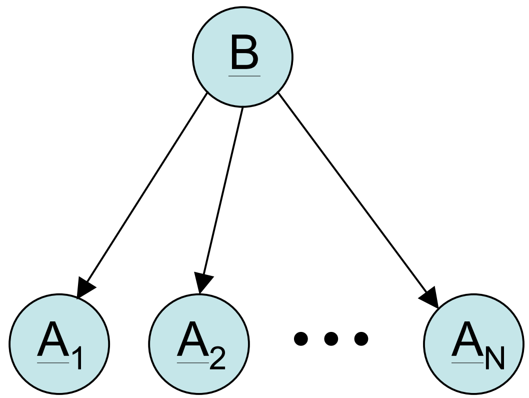
external representation in terms of aggregate properties



internal representation in terms of component properties

Rimey 93:

Tree-shaped part-of nets, is-a trees, expected-area nets, and task nets



unrealistic conditional independence:

$$P(A_1 \ldots A_N | B) = P(A_1|B)\, P(A_2|B) \ldots P(A_N|B)$$

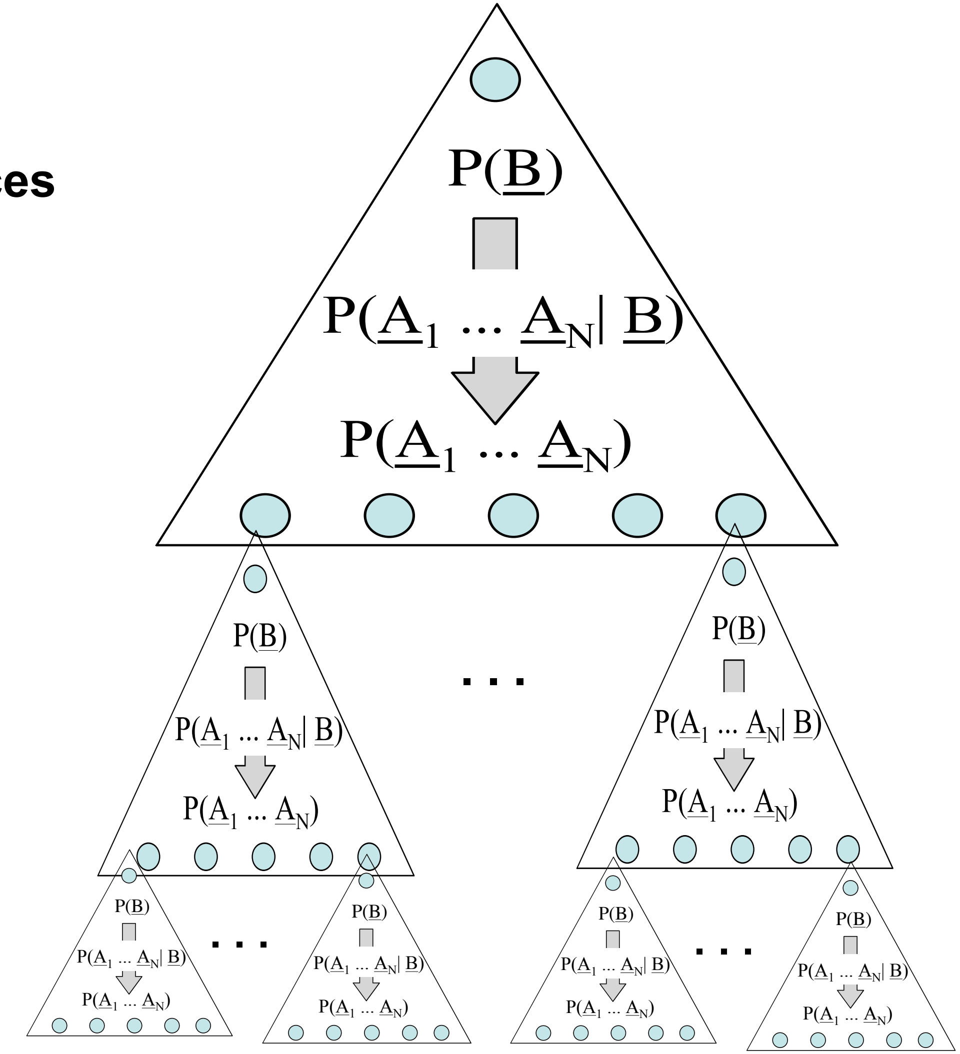
Probabilistic Aggregate Hierarchy

What are useful (and plausible) independence assumptions

- **for efficient probabilistic inferences**
- **for intuitive aggregate models?**

Simplifying assumptions (initially):

- **Distinct names for multiple parts of the same kind**
- **Fixed set of parts per aggregate**
- **No specialization branchings**



Bayesian Compositional Hierarchy (1)

Conditional-independence requirements for a compositional hierarchy to be an "Bayesian compositional hierarchy":

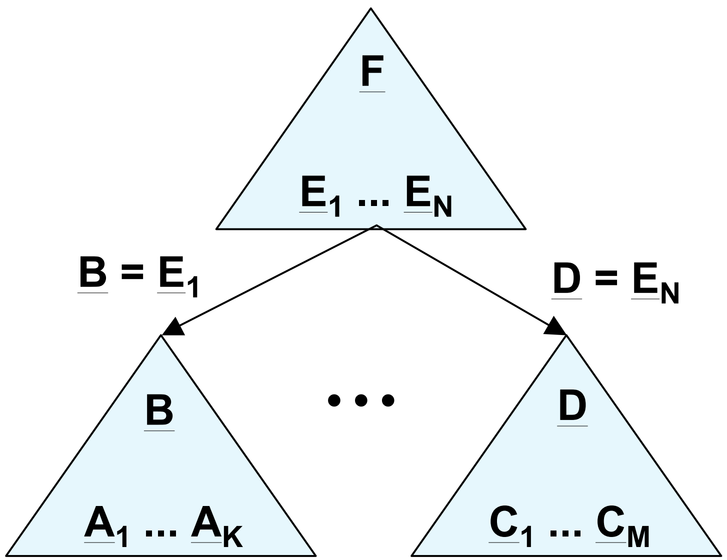
X	**an aggregate node**
Y_1 .. Y_N	**the parts of X**
succ(X)	**all successors of X**

Req 1: $P(X \mid succ(X)) = P(X \mid Y_1 .. Y_N)$ **(1)**

Aggregate properties do not depend on details below the part properties.

Example:

Given E_1 ... E_N, F is independent of all successors below E_1 ... E_N



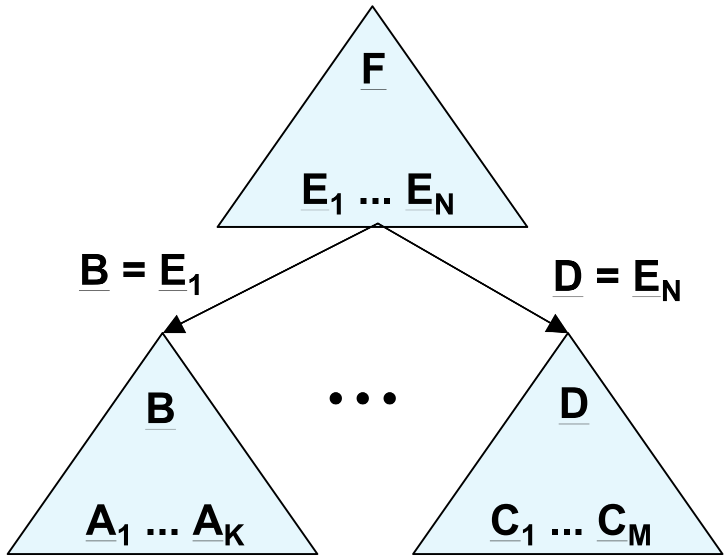
Bayesian Compositional Hierarchy (2)

Req 2: $P(\, succ(\underline{Y}_i) \mid \underline{Y}_1 \,..\, \underline{Y}_N \,) = P(\, succ(\underline{Y}_i) \mid \underline{Y}_i \,)$ **(2)**

Part properties depend only on the properties of the corresponding mother aggregate.

Example:

Given $\underline{B} = \underline{E}_1$,
$\underline{A}_1 \ldots \underline{A}_K$ and their successors are independent of $\underline{E}_2 \ldots \underline{E}_N$



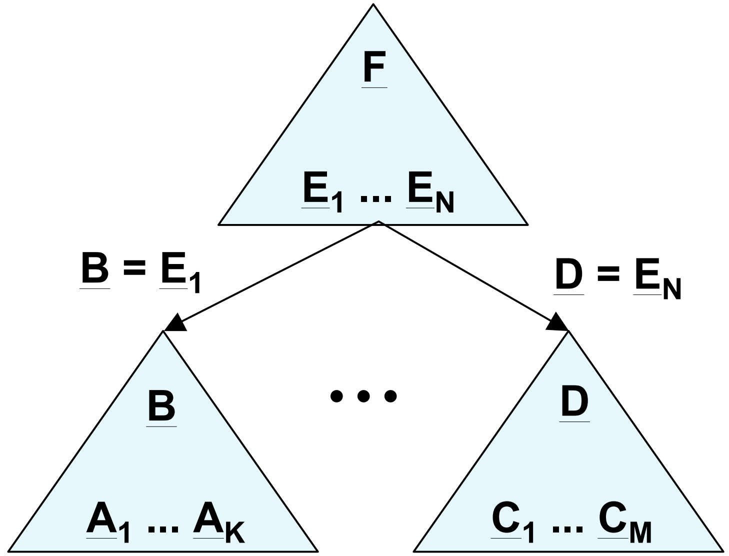
Bayesian Compositional Hierarchy (3)

Req 3: $P(\,succ(\underline{Y}_1 .. \underline{Y}_N) \mid \underline{Y}_1 .. \underline{Y}_N\,) = \prod P(\,succ(\underline{Y}_i) \mid \underline{Y}_1 .. \underline{Y}_N\,)$ **(3)**

Parts of different aggregates are statistically independent given their mother aggregates.

Example:

Given $\underline{E}_1 \ldots \underline{E}_N$,
$\underline{A}_1 \ldots \underline{A}_K$ and their successors are independent of $\underline{C}_1 \ldots \underline{C}_M$ and their successors



From (2) and (3) it follows that

$$P(\,succ(\underline{Y}_1 .. \underline{Y}_N) \mid \underline{Y}_1 .. \underline{Y}_N\,) = \prod P(\,succ(\underline{Y}_i) \mid \underline{Y}_i\,)$$

Bayesian Compositional Hierarchy (4)

$$P(\text{all}) = P(\underline{X} \mid succ(\underline{X}))\ P(succ(\underline{X}))$$

$$= P(\underline{X} \mid \underline{Y}_1 \,..\, \underline{Y}_N)\ P(succ(\underline{X})) \qquad \text{by Req 1}$$

$$= P(\underline{X} \mid \underline{Y}_1 \,..\, \underline{Y}_N)\ \ P(\ \underline{Y}_1 \,..\, \underline{Y}_N\ succ(\ \underline{Y}_1 \,..\, \underline{Y}_N))$$

$$= P(\underline{X} \mid \underline{Y}_1 \,..\, \underline{Y}_N)\ \ P(succ(\ \underline{Y}_1 \,..\, \underline{Y}_N) \mid \underline{Y}_1 \,..\, \underline{Y}_N\ )\ \ P(\underline{Y}_1 \,..\, \underline{Y}_N\ )$$

$$= P(\underline{X} \mid \underline{Y}_1 \,..\, \underline{Y}_N)\ \ \prod P(succ(\ \underline{Y}_i) \mid \underline{Y}_i\ )\ \ P(\underline{Y}_1 \,..\, \underline{Y}_N) \qquad \text{by Req 2 + 3}$$

$$\Rightarrow\ P(succ(\underline{X}) \mid \underline{X}) = P(\underline{Y}_1 \,..\, \underline{Y}_N \mid \underline{X})\ \ \prod P(succ(\ \underline{Y}_i) \mid \underline{Y}_i\ )$$

Recursive application gives:

$$P(\ \underline{Z}_0 \,..\, \underline{Z}_M\ ) = P(\underline{Z}_0) \prod_{i = 0 \,...\, M} P(\ parts(\underline{Z}_i) \mid \underline{Z}_i\ )$$

Z_0 is a node and Z_i, $i = 1 \,..\, M$ are its successors.

The complete JPD of an abstraction hierarchy can be computed from the conditional aggregate JPDs.

Probability changes may be propagated along tree-shaped hierarchy.

Alternative Formalization of Bayesian Compositional Hierarchy

External properties $\underline{Z}$ of an aggregate are determined by the functional mapping f: parts($\underline{Z}$) => $\underline{Z}$

$\Rightarrow$ P($\underline{Z}$ | parts($\underline{Z}$)) is known and fixed

The Bayesian Compositional Hierarchy factorization formula can be reformulated:

$$P(\underline{Z}_0 .. \underline{Z}_M) = \Pi\ P(\underline{Z}_i \mid parts(\underline{Z}_i))\ C(parts(\underline{Z}_i))$$

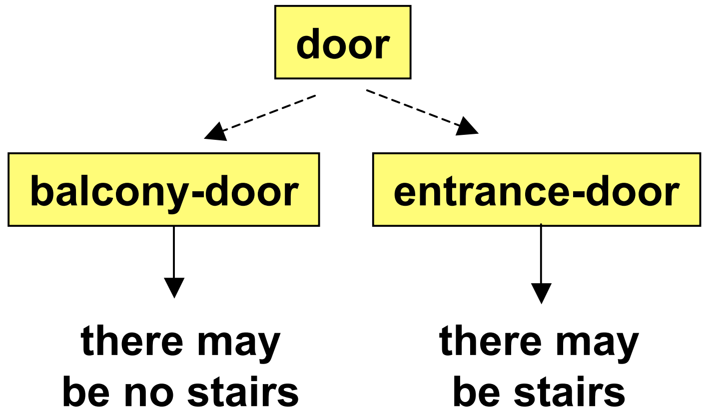
$$\text{where } C(\underline{Y}_1 .. \underline{Y}_N) = P(\underline{Y}_1 .. \underline{Y}_N) / \Pi\ P(\underline{Y}_i)$$

Given the probability distributions of the properties of individual parts, one can construct a hierarchy <u>bottom-up</u> by determining the correlations between parts belonging to an aggregate.

$\Rightarrow$ Unsupervised learning of aggregates

Choice of Alternative Specializations

Specializing a hypothesis

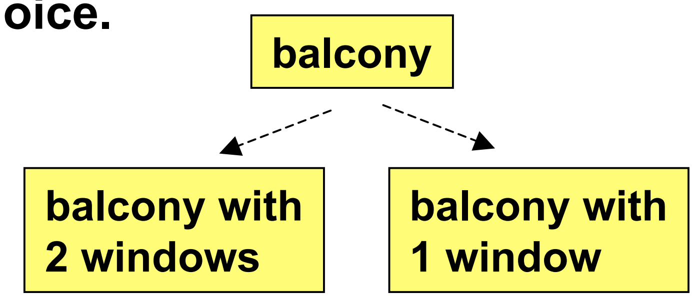


Disjunctive specializations can be modelled probabilistically, probability changes of one disjunctive branch may be propagated to the other branch.

Evidence assignment to one disjunctive branch forces specialization decision and must prohibit evidence assignment to the other branch.

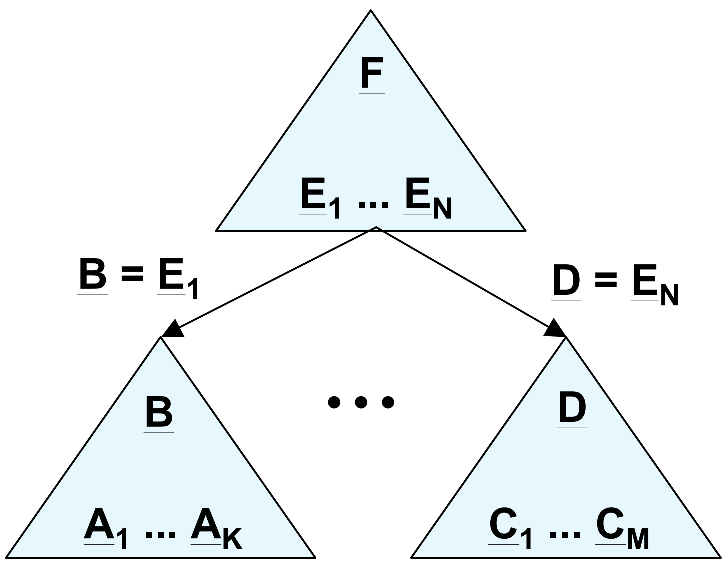
Currently, specialization decisions in SCENIC may be taken top-down, causing backtracking in the case of a wrong choice.

Aggregates with different cardinalities may be modelled as disjunctive specializations => the same applies.

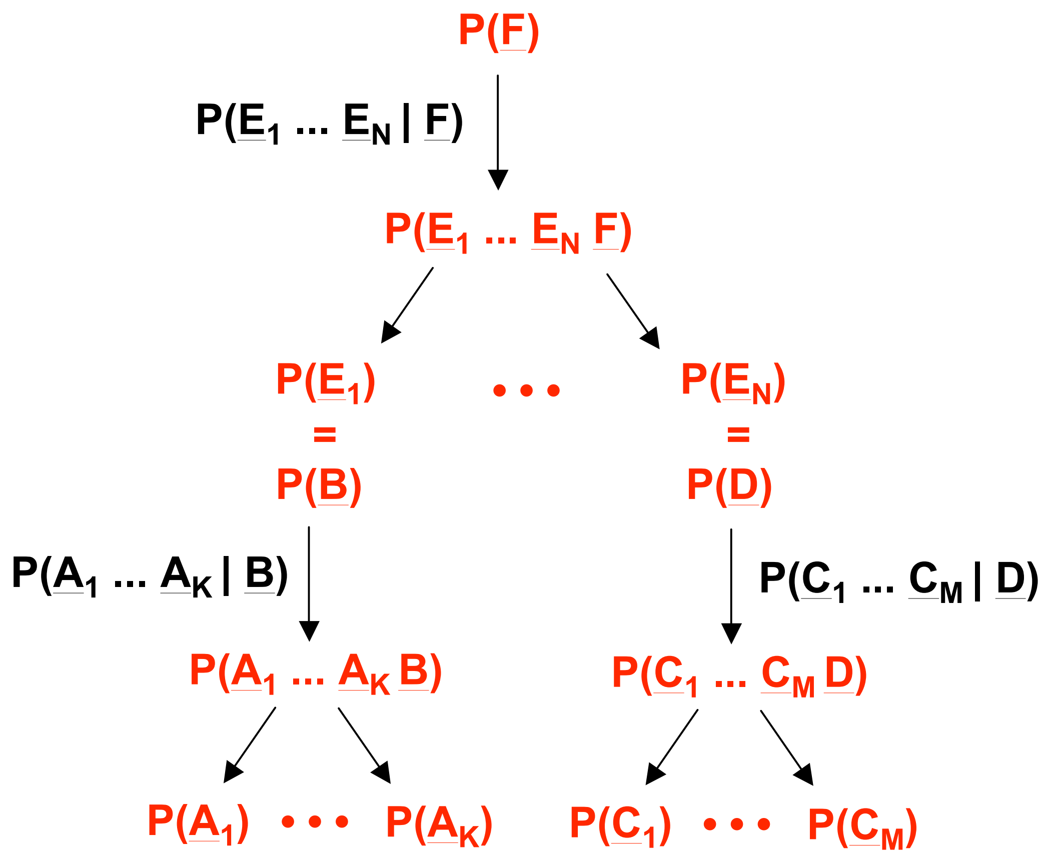


Top-down Initialization

Aggregate hierarchy subtree:



Sequence of computations:

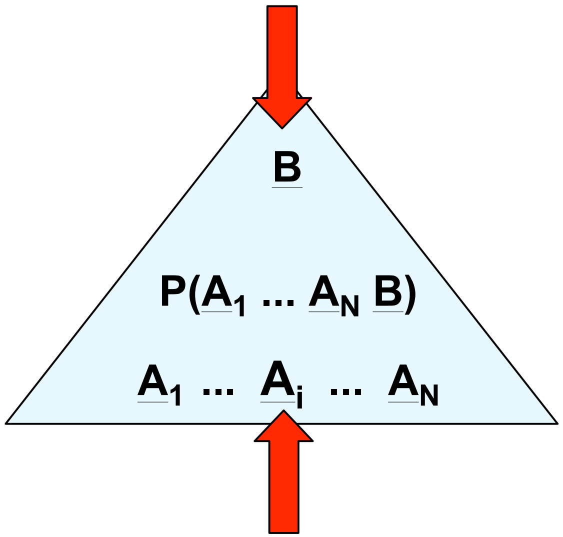


Change Propagation

After initialization, the state of each aggregate is represented by $P(\underline{A}_1 \dots \underline{A}_N)$ with marginalizations $P(\underline{A}_i)$, $i = 1 \dots N$, and $P(\underline{B})$.

A change has to be propagated if $P(\underline{B}) \Rightarrow P'(\underline{B})$ or $P(\underline{A}_i) \Rightarrow P'(\underline{A}_i)$, some i.

Crisp evidence $\underline{e}$ for $\underline{A}_i$ is modelled as $P(\underline{A}_i = \underline{e}) = 1$ and $P(\underline{A}_i \neq \underline{e}) = 0$.



Propagating down:

$P(\underline{B}) \Rightarrow P'(\underline{B})$

$P'(\underline{A}_1 \dots \underline{A}_N \underline{B}) = P(\underline{A}_1 \dots \underline{A}_N \underline{B})\, P'(\underline{B}) / P(\underline{B})$

followed by marginalizations

Propagating up:

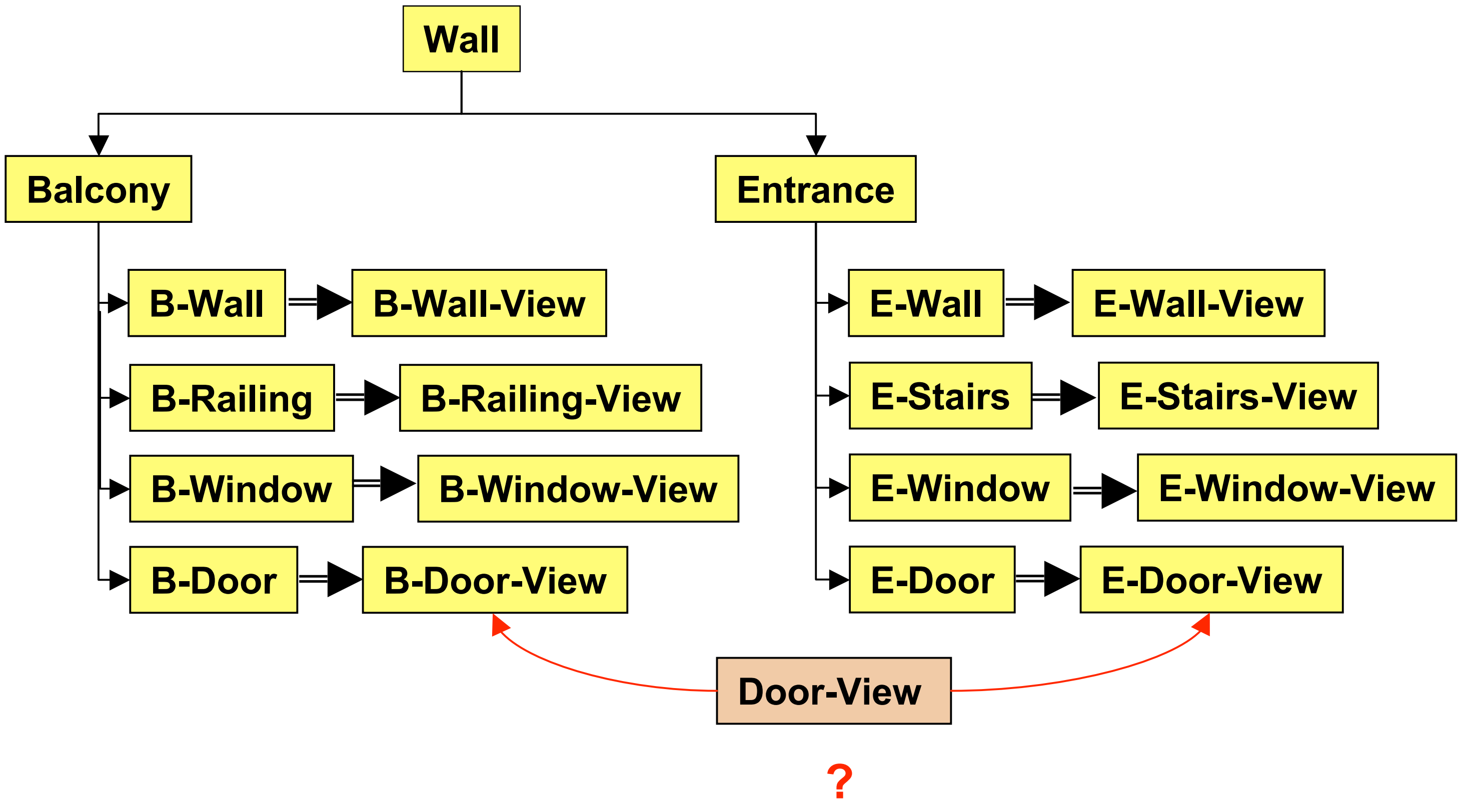
$P(\underline{A}_i) \Rightarrow P'(\underline{A}_i)$

$P'(\underline{A}_1 \dots \underline{A}_N \underline{B}) = P(\underline{A}_1 \dots \underline{A}_N \underline{B})\, P'(\underline{A}_i) / P(\underline{A}_i)$

followed by marginalizations

Preference Computation for Evidence Classification

- Probabilities within a branch may be compared without considering the rest of the compositional hierarchy
- Probabilities are updated after each decision and influence the following decisions



Best-first Evidence Classification

Stepwise procedure

A Choose evidence which allows most certain classification (reducing need for backtracking)

$$\text{all } i \neq k: P(view_k \mid e) >> P(view_i \mid e)$$

B If there is no probable classification for a given piece of evidence,
- perform backtracking to revise previous classifications, or
- request low-level validation of evidence

C Determine revised $P(view_i \mid e_j)$ after each classification
=> evidence propagation in probabilistic hierarchy

D Repeat steps A - D until task is completed
- evidence is exhausted
- scene interpretation is sufficiently certain
- specific interpretation request can be answered
- no conceptual model fits evidence

How to Determine Probability Distributions for Aggregates

1. **Determine distributions for known crisp aggregates**

   Two alternative approaches:

   a. **Determine JPDs of internal and external properties by statistics (frequentist approach).**
   b. **Estimate JPDs based on human experiences and the mappings from internal to external properties.**

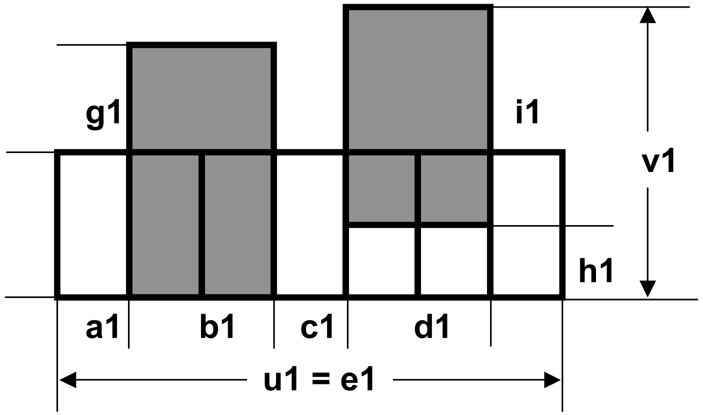
2. **Learn aggregate concepts from scratch**

   - **Observe primitives, determine statistics**
   - **Build aggregate hierarchy by agglomerative clustering (use distance measure to establish Bayesian abstraction)**
   - **Derive higher-level probabilities from lower-level probabilities**

Gaussian Aggregate Models

Uncertain aggregate properties can sometimes be roughly modelled as Gaussian densities.

Example:



Balcony probability densities:

$p_{b\text{-}door}(b1\ g1)$

$p_{b\text{-}window}(d1\ i1)$

$p_{railing}(b1\ g1)$

$p_{balcony\text{-}int}(a1\ b1\ c1\ d1\ e1\ f1\ g1\ h1\ i1)$

$p_{balcony\text{-}ext}(u1\ v1)$

$u1 = e1$

$v1 = h1 + i1$

} **must be linear combination of parts properties**

Probabilistic representation of the aggregate "balcony" by

$$P(\underbrace{a1\ b1\ c1\ d1\ e1\ f1\ g1\ h1\ i1}_{\text{parts properties}} \mid \underbrace{u1\ v1}_{\text{external aggregate properties}})$$

Probabilistic Balcony Description

Specification of $N(\mu, \Sigma)$ for balcony properties by human estimates (unit = 1 dcm)

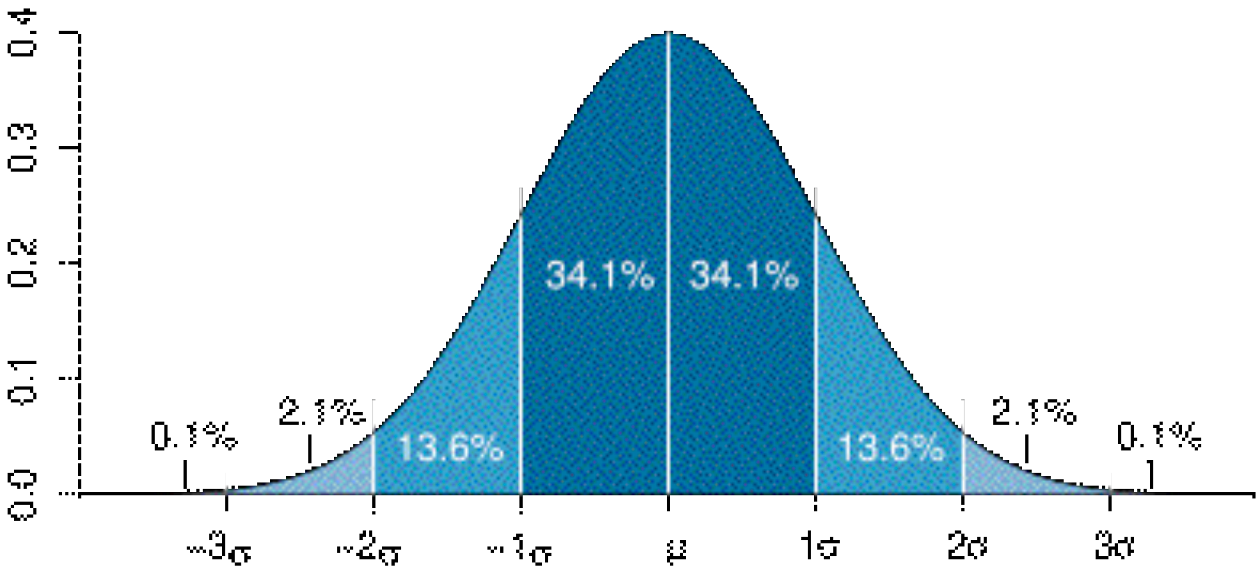
Means

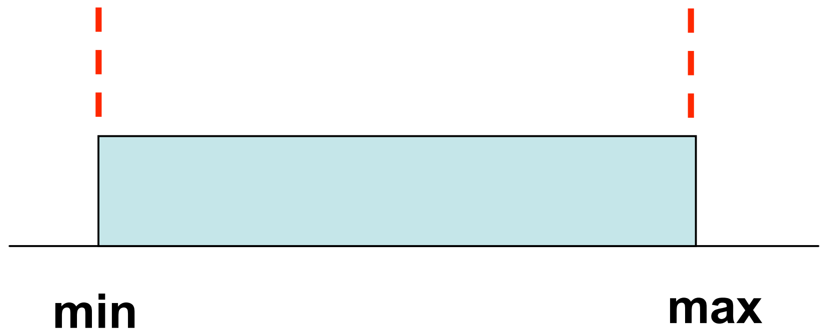
a1	b1	c1	d1	e1	f1	g1	h1	i1	u1	v1
5	9	5	15	39	12	19	12	15	39	27

Covariances

	a1	b1	c1	d1	e1	f1	g1	h1	i1	u1	v1
a1	6,0	1,2	3,3	6,0	3,5	0,0	0,0	0,0	0,0	3,5	0,0
b1	1,2	2,3	1,2	5,3	2,1	0,0	0,4	0,0	1,2	2,1	1,2
c1	3,3	1,2	6,0	6,0	3,5	0,0	0,0	0,0	0,0	3,5	0,0
d1	6,0	5,3	6,0	60,0	11,0	0,0	0,0	0,0	8,5	11,0	8,5
e1	3,5	2,1	3,5	11,0	20,0	0,0	0,0	0,0	0,0	20,0	0,0
f1	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0
g1	0,0	0,4	0,0	0,0	0,0	0,0	0,3	0,0	0,4	0,0	0,4
h1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	2,3	0,0	0,0	2,3
i1	0,0	1,2	0,0	8,5	0,0	0,0	0,4	0,0	6,0	0,0	6,0
u1	3,5	2,1	3,5	11,0	20,0	0,0	0,0	0,0	0,0	20,0	0,0
V1	0,0	1,2	0,0	8,5	0,0	0,0	0,4	2,3	6,0	0,0	8,3

Normal Distributions vs. SCENIC Ranges

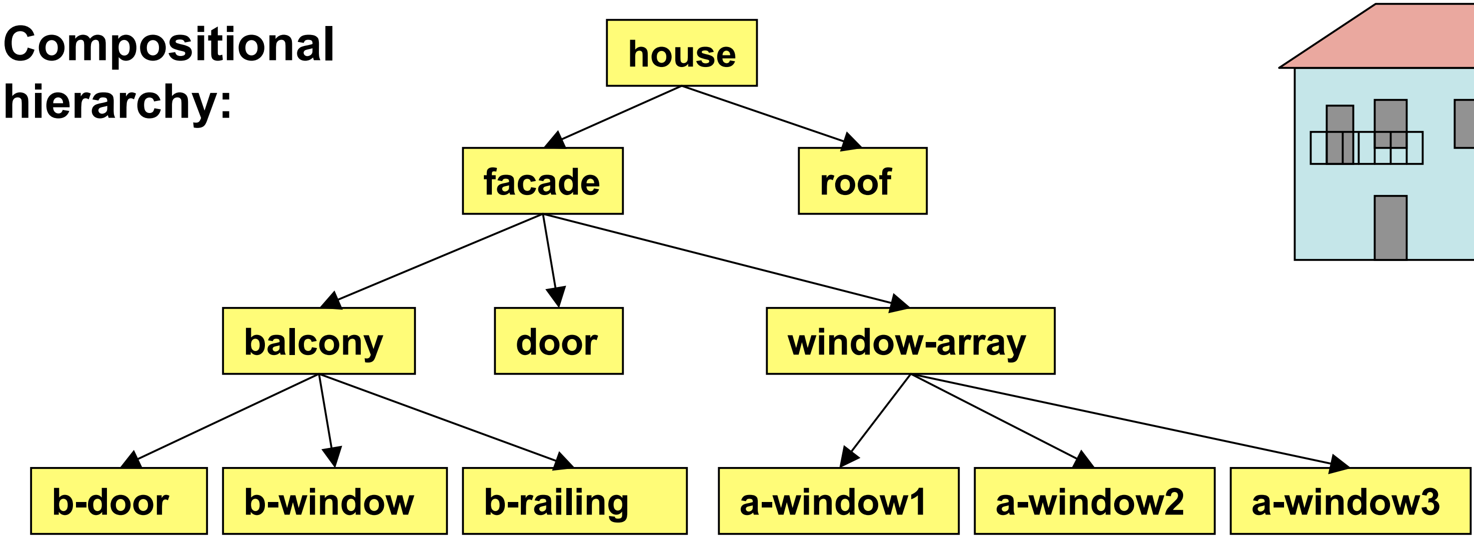


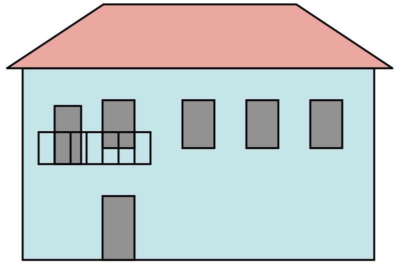


- **Gaussian range [-2σ .. 2σ] may be used for SCENIC variables with range type values**
- **Exploitation restricted to values in this range**

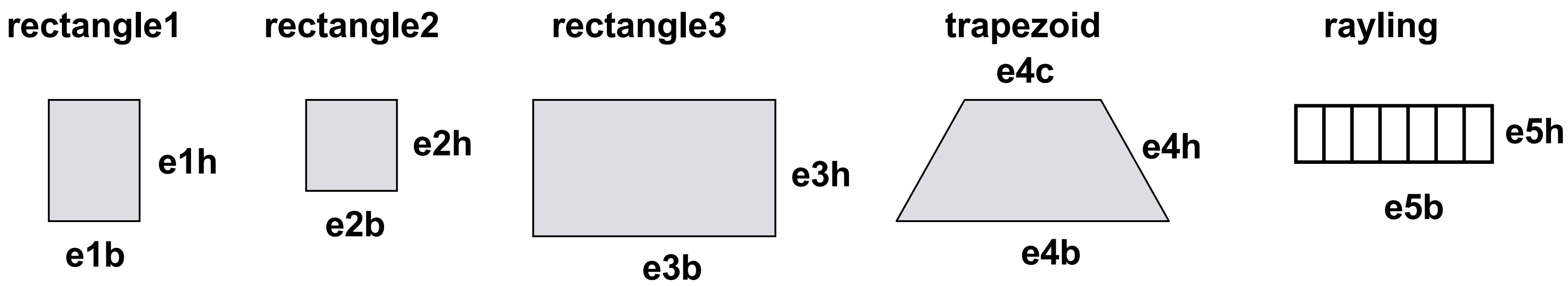
"Major" Example for a Gaussian Aggregate Hierarchy

Compositional hierarchy:



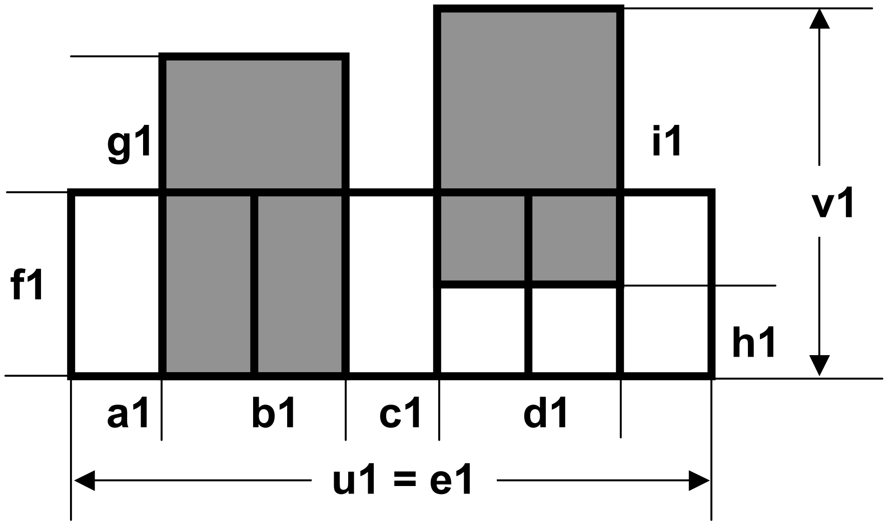


Evidence:



+ absolute image positions

Variables for Balcony and Window-Array



Balcony probability densities:

$p_{b\text{-}door}(b1\ g1)$

$p_{b\text{-}window}(d1\ i1)$

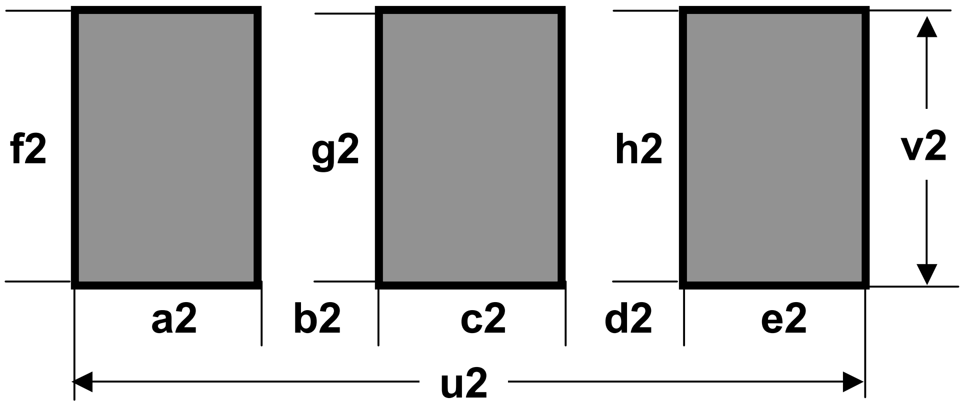
$p_{railing}(b1\ g1)$

$p_{balcony\text{-}int}(a1\ b1\ c1\ d1\ e1\ f1\ g1\ h1\ i1)$

$p_{balcony\text{-}ext}(u1\ v1)$

$u1 = e1$

$v1 = h1 + i1$



Window-array probability densities:

$p_{a\text{-}window1}(a2\ f2)$

$p_{a\text{-}window2}(c2\ g2)$

$p_{a\text{-}window3}(e2\ h2)$

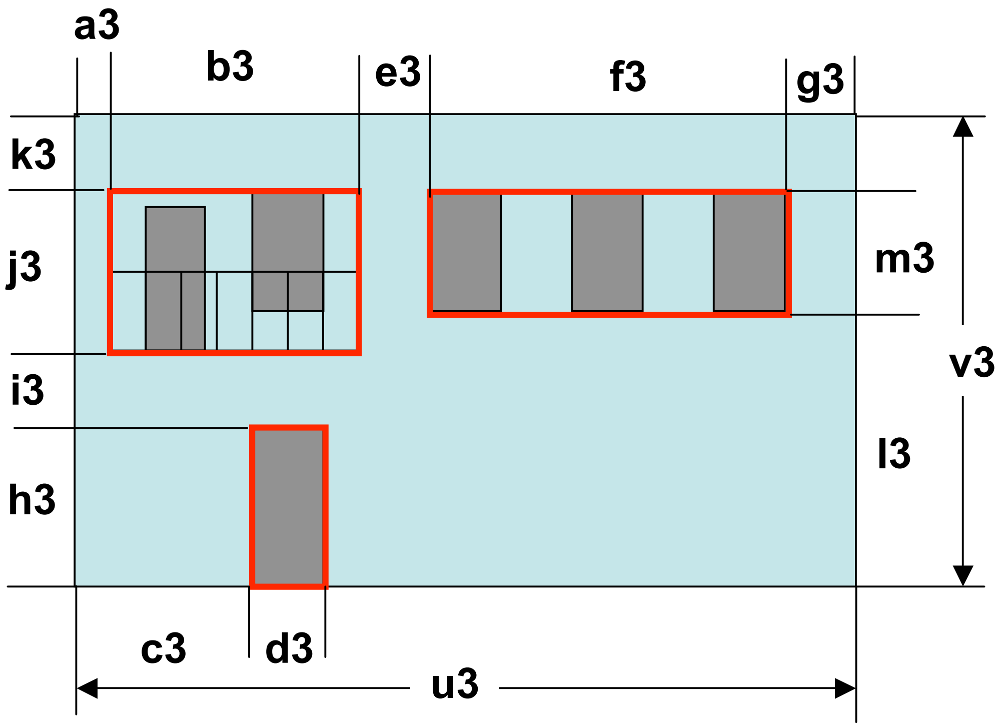
$p_{window\text{-}array\text{-}int}(a2\ b2\ c2\ d2\ e2\ f2\ g2\ h2)$

$p_{window\text{-}array\text{-}ext}(u2\ v2)$

$u2 = a2 + b2 + c2 + d2 + e2$

$v2 = (f2 + g2 + h2)/3$

Variables for Facade



Facade probability densities:

$p_{door}(d3\ h3)$

$p_{balcony}(b3\ j3)$

$p_{window\text{-}array}(f3\ m3)$

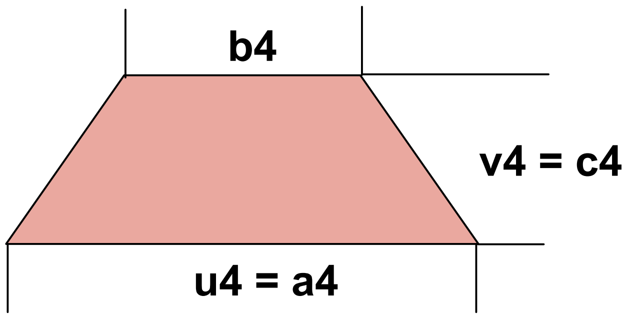
$p_{facade\text{-}int}(a3\ b3\ c3\ d3\ e3\ f3\ g3\ h3\ i3\ j3\ k3\ l3\ m3)$

$p_{facade\text{-}ext}(u3\ v3)$

$u3 = a3 + b3 + e3 + f3 + g3$

$v3 = h3 + i3 + j3 + k3$

Variables for Roof and House



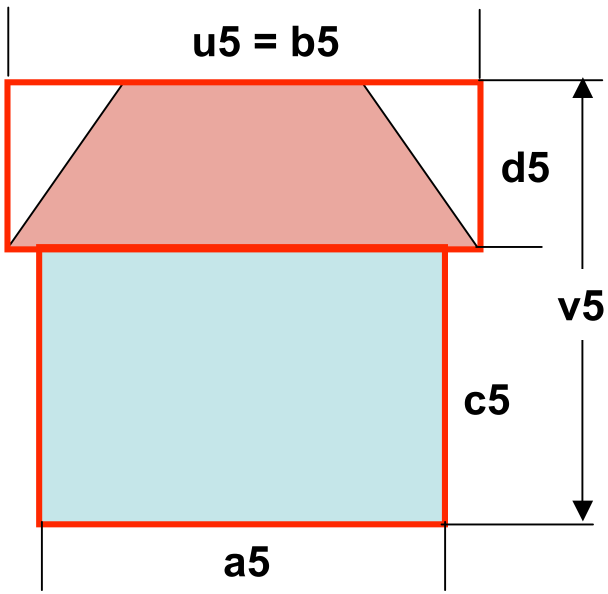
Roof probability densities:

$P_{roof\text{-}int}(a4\ b4\ c4)$

$P_{roof\text{-}ext}(u4\ v4)$

u4 = a4

v4 = c4



House probability densities:

$P_{facade}(a5\ c5)$

$P_{roof}(b5\ d5)$

$P_{house\text{-}int}(a5\ b5\ c5\ d5)$

$P_{house\text{-}ext}(u5\ v5)$

u5 = b5

v5 = c5 + d5

Summary and Perspectives

- Scene interpretation can be viewed as partial model construction (with multiple solutions)

- Probabilities provide
  - a global preference measure for choosing between alternative interpretations
  - a local preference measure for choosing promising stepwise decisions

- Bayesian Compositional Hierarchies provide plausible abstractions and allow efficient probability propagation

- To be explored:
  - how to deal with optional parts
  - how to deal with disjunctive choices
  - how to restrict propagations to areas of interest
  - how to learn probability distributions

  . . .