

# Spatial Cognition for Mobile Robots: A Hierarchical Probabilistic Concept- Oriented Representation of Space

Shrihari Vasudevan

Advisor: Prof. Dr. Roland Siegwart

Autonomous Systems Lab, ETH Zurich, Switzerland.

[shrihari.vasudevan@ieee.org](mailto:shrihari.vasudevan@ieee.org)



# Outline

- Motivation of work
  - Vision & Issues with SOA
- Approach
- From Objects to Places
  - Perception – Representation – Cognition
- From Objects to **Concepts** to Places
  - Overview of different Conceptualization models
  - Object count + relationship model (M4) - specifics
- User Study
- Applications and Contributions of work

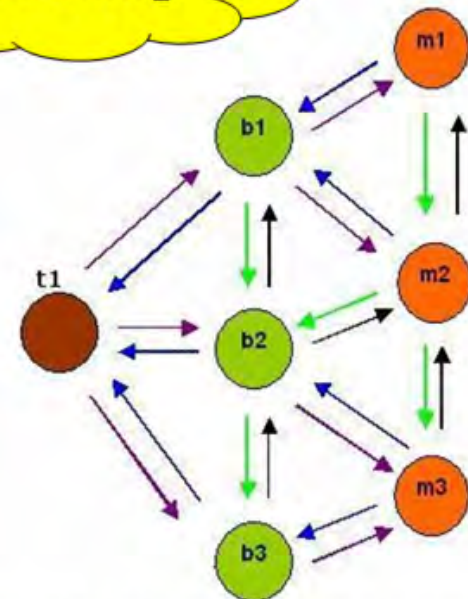
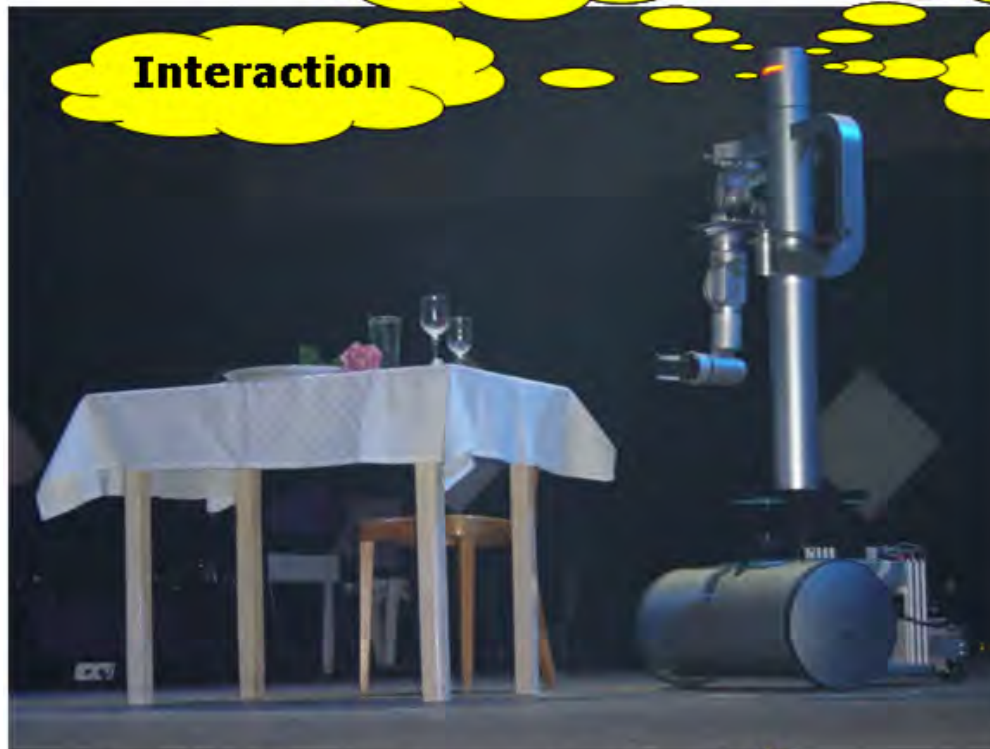
## Cognitive Robot Companions of the future

Navigation

Manipulation

Interaction

Reasoning



Spatial Relationships between objects, regions and places

Figure - Robot in a typical home scenario

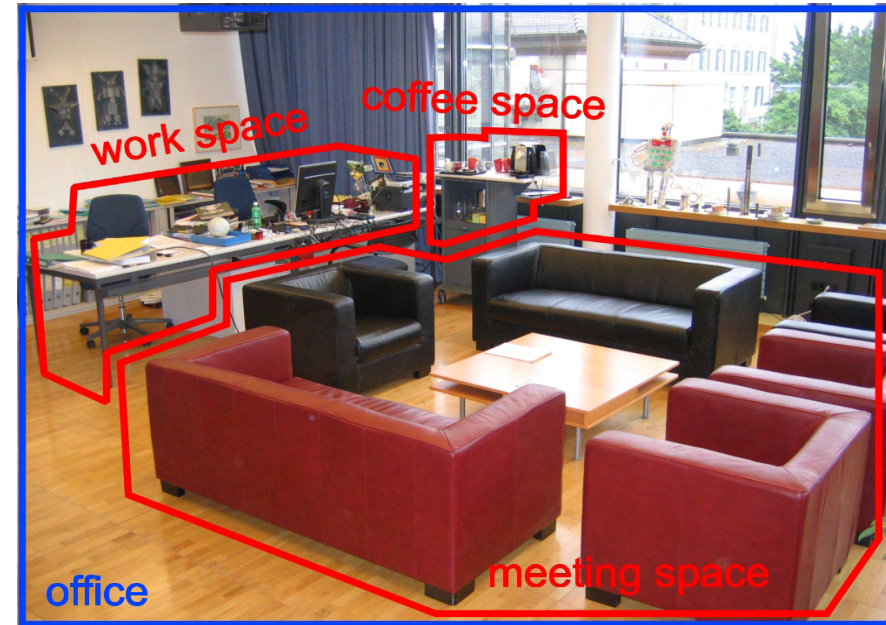
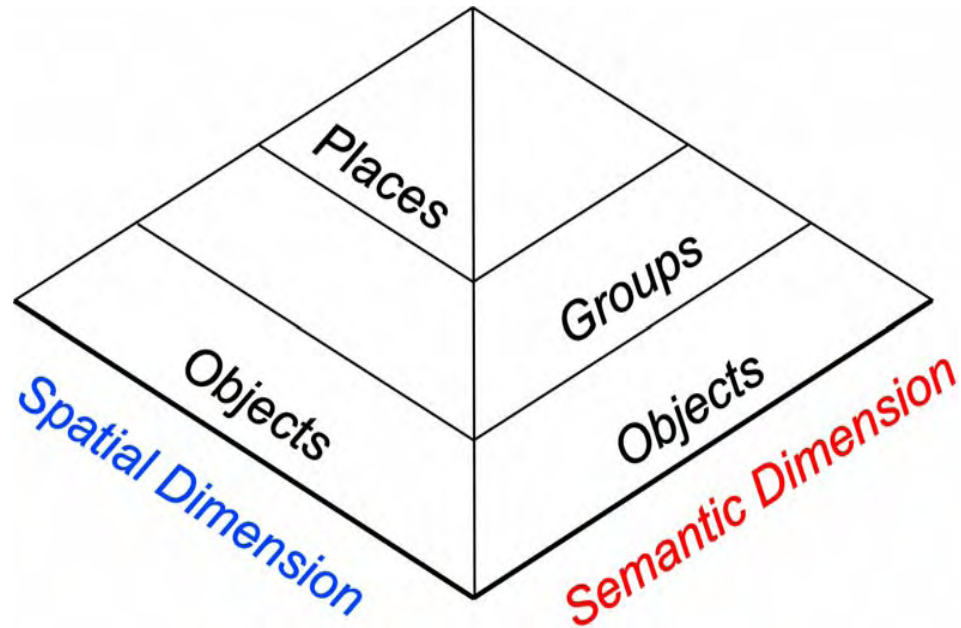
**My Work**

**REPRESENTATION FOR ROBOT**  
Hierarchical Probabilistic Representation  
of Space

# Motivation

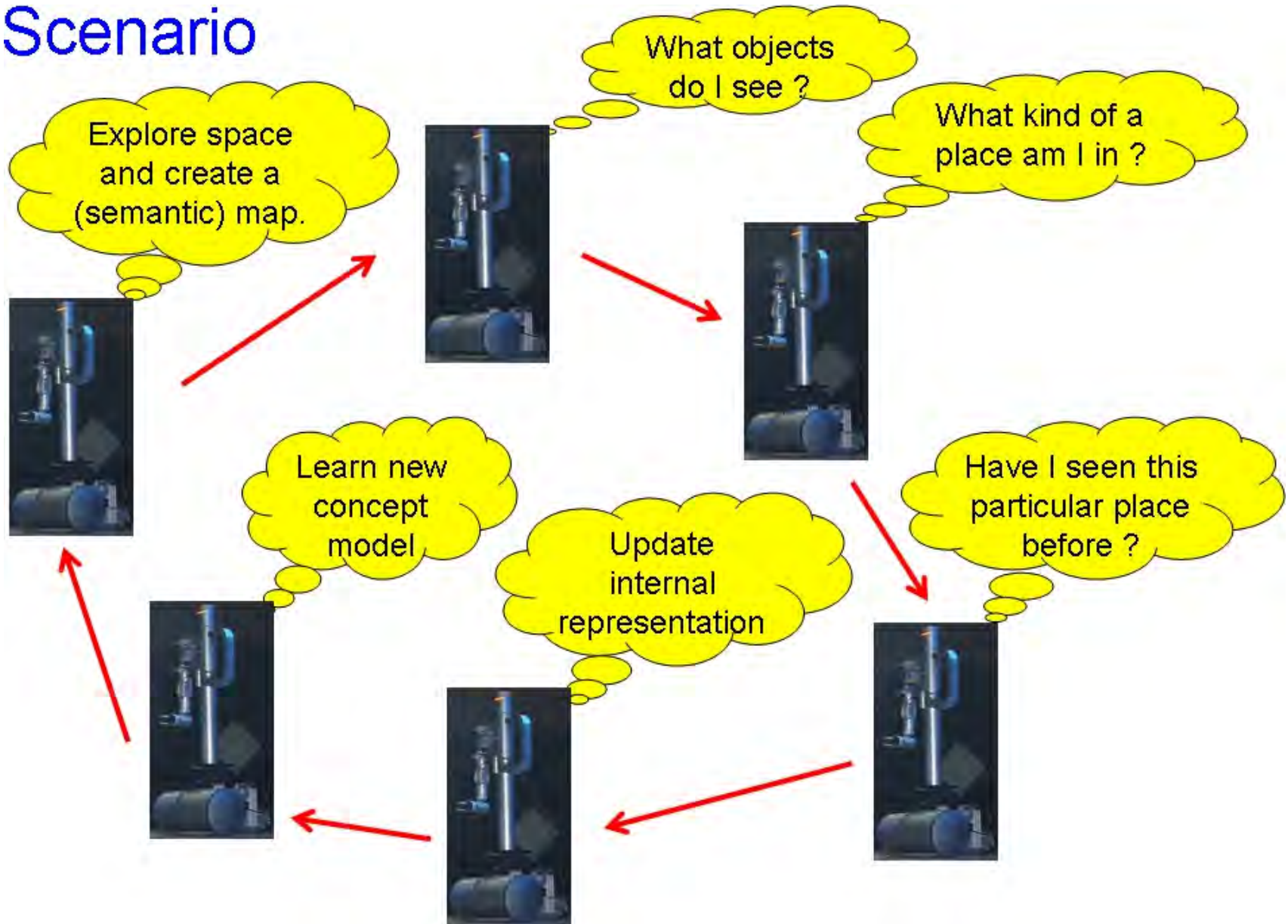
- Robot mapping – well researched problem !
  - Metric maps
  - Topological maps
  - Hybrid maps
- Issues
  - Only suited to robot navigation
  - Doesn't encode much or most of the **semantics**
  - **Spatial awareness** of robots – modest.

# Approach



- **Enhance feature set**  
high level features - objects, doors etc.
- **Graph based representation**  
objects and relationships between them
- **Abstraction towards hierarchy**  
from sensory information to increasingly abstract concepts (spatial / semantic).
- **Reference: Vasudevan@LNAI07**
- Semantic abstraction (groups) → Objects are clustered into groups to capture their semantics (functionality / arrangement etc.)
- Spatial abstraction (places) → A collection of groups of objects, typically formed by structural/ boundary elements such as doors and walls.

# Scenario



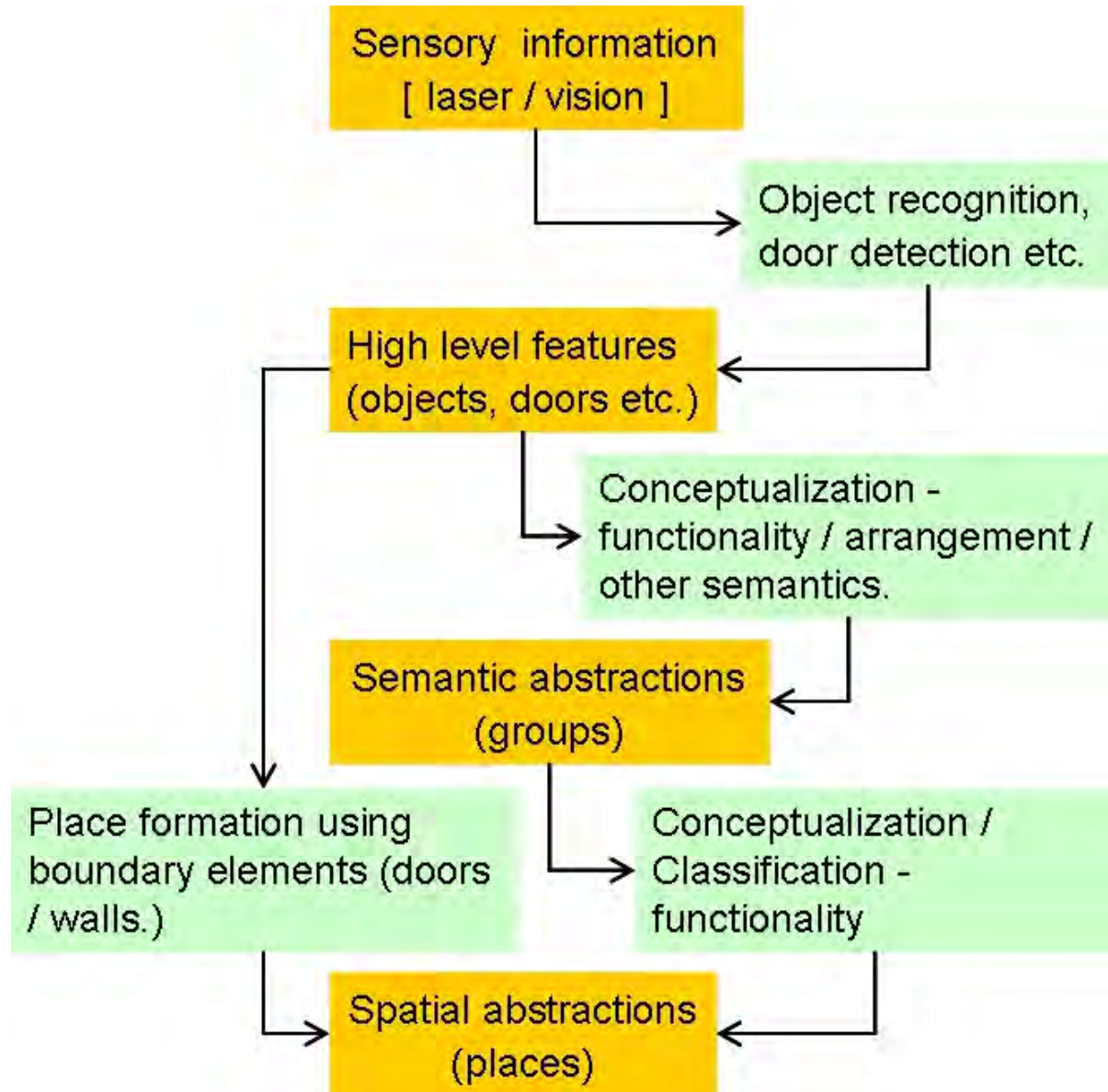
# Outcome: Office

## (movie of experiment on robot platform)



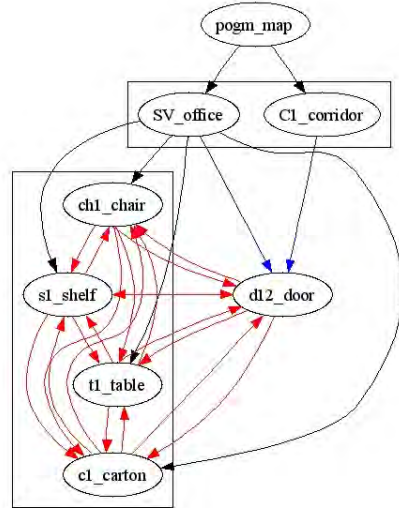
- Movie demonstrating thesis title
  - Probabilistic Object graph representation of space
  - Bayesian Conceptualization & Place Classification
- Movie shows perception, representation / mapping and cognition in an integrated demonstration.
  - inference using a-priori learnt models from a dataset

# Implementation Overview

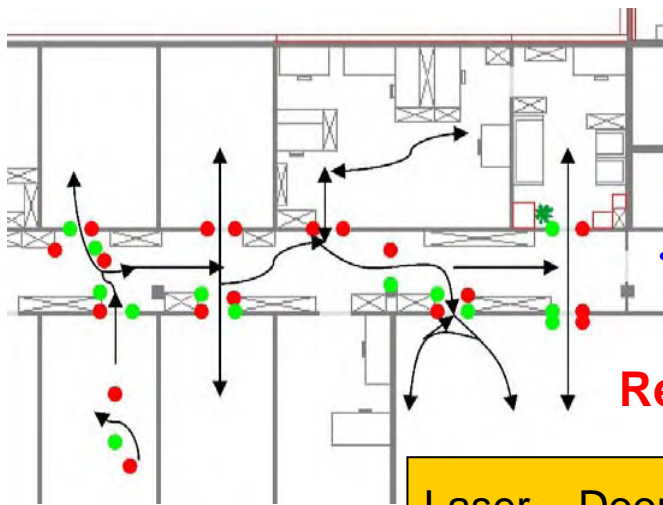


# POGM

## Robot exploration / mapping

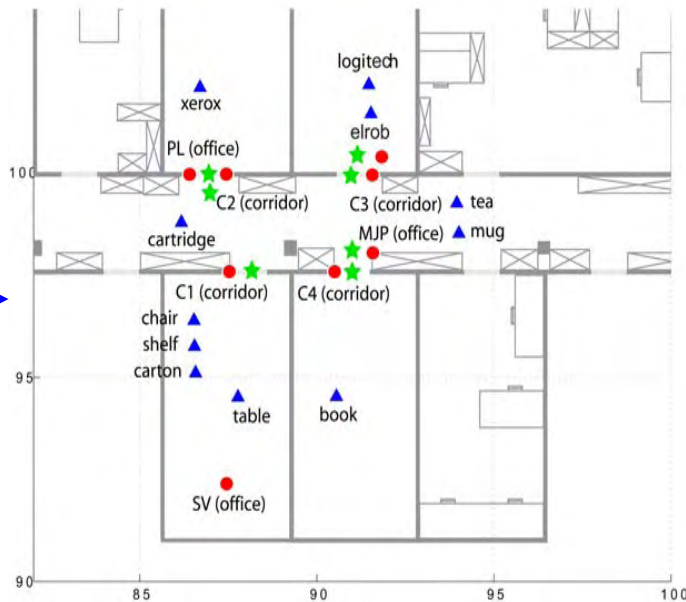


## Stereo - SIFT Object recognition



## Laser - Door detection

Ref: Vasudevan@RAS07



## Object based map

# Probabilistic graphical representation

- Representation encodes objects and relationships between them.
- Relationships are encoded in terms of relative spatial information (distance and angle measures in 3D space) between objects. These were meant to metrically capture typical spatial relationships such as nearness, to-the-right-of, above, below etc.
- Both existential (discrete belief measure) and position beliefs (Gaussian - mean / covariance matrix) are maintained and treated separately.
- Design = extended (over several places) relative metric map or global topological map of local metric maps.

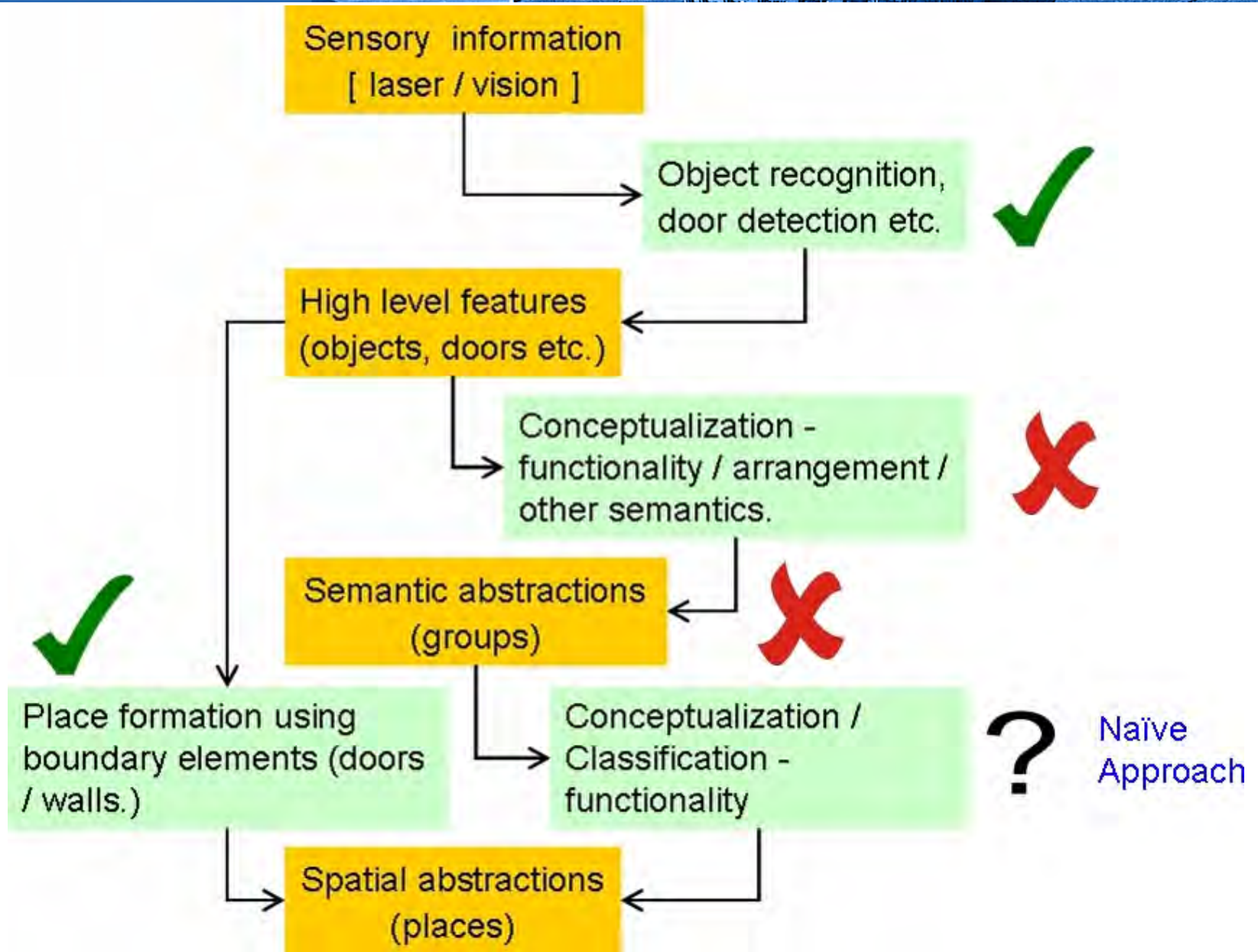
# Place Classification (M1)

$$P(\text{object\_type} | \text{place\_type}) = \frac{N_{\text{object\_type}} + 1}{N_{\text{place\_type}} + 2}$$

## Algorithm

1. Initialize each place\_type with a prior belief.
2. For each scene that the robot observes
  - For each object\_type 'o' that is observed in unknown place
    - (a) For each place\_type 'p' that could have this particular object, compute posterior belief of place\_type given object\_type using Bayes rule
$$P(p|o) = \frac{P(o|p)*P(p)}{P(o)}$$
  - Normalize beliefs of places
  - Hypothesis selection based on distinctiveness and threshold.

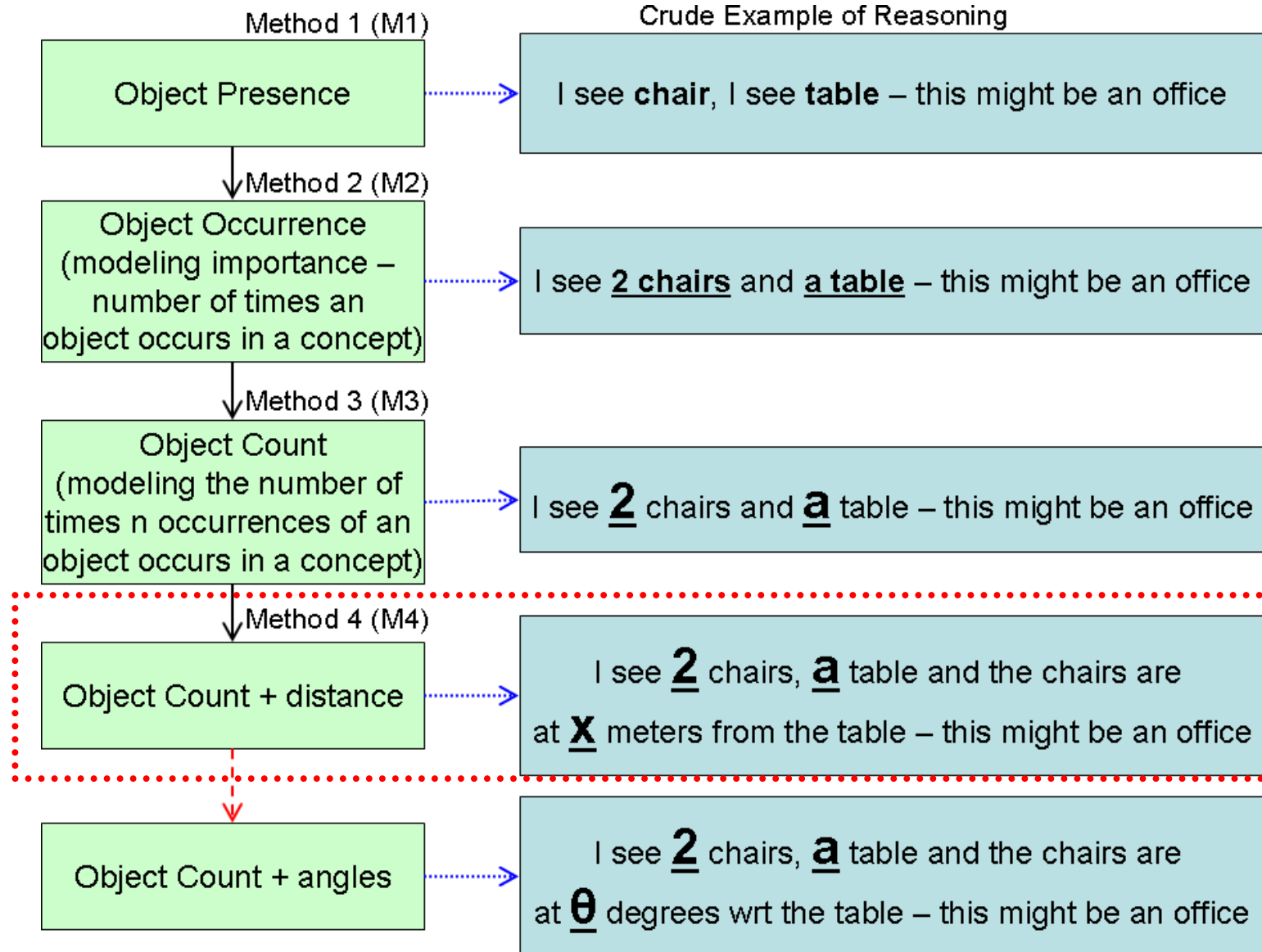
Spatial Cognition – Place classification & Place recognition (graph match).



# Issues at this point

- Semantics only represented by the presence of objects.
- Place classification was naïve (M1)
  - No learning from negative exemplars
  - Inference based only on evidence present and not on that which was absent.
  - Did not deal with multi-object occurrences and only worked on the basis of object category presence/absence.
- Thus, motivation for the work that follows –
  - **Improve place classification** procedure
  - **Increase semantic content** in robot maps. (i.e. form **concepts**)

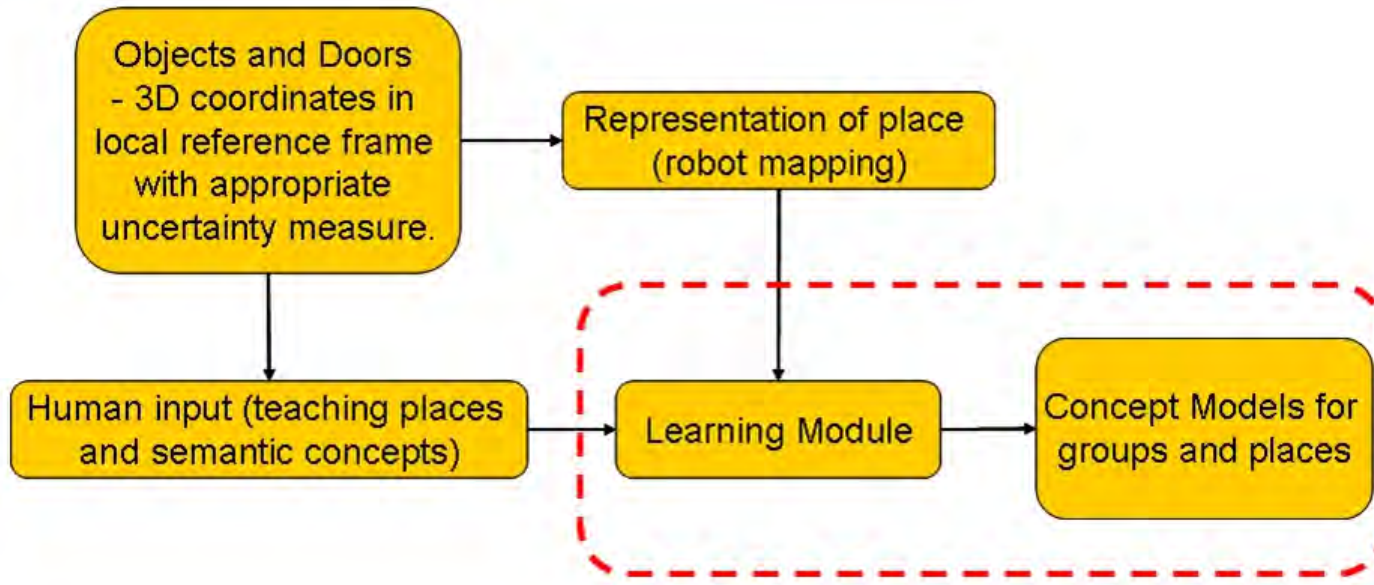
# Overview of attempted approaches



Reference:  
Vasudevan  
@RAS08

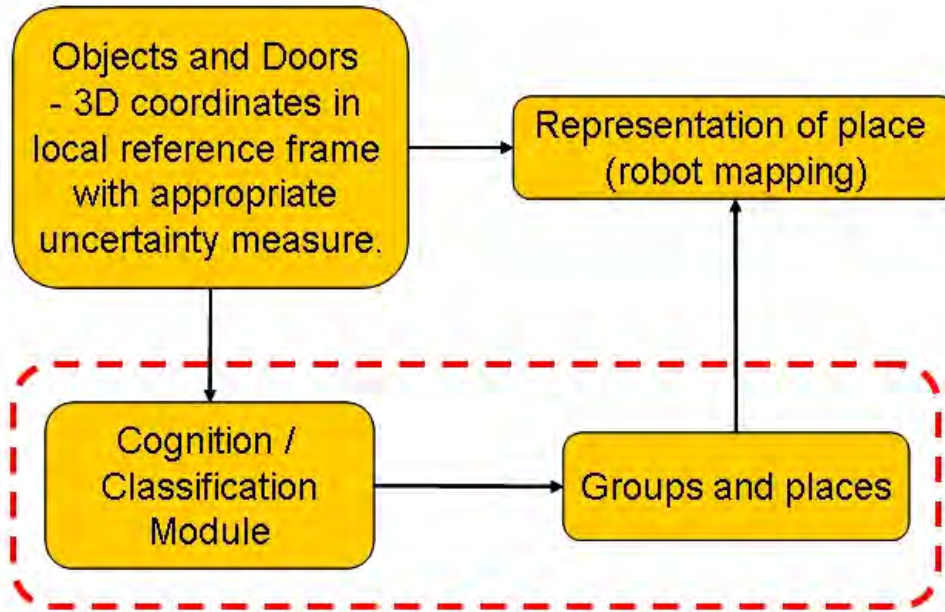
Preliminary  
Results for  
OC + angles

# General Approach: Learning



- No specific ontology used.
- Robot has an object recognition capability.
- Human user “teaches” (or bootstrap with a database of exemplars) the robot different concept instances. (Home Tour scenario).

# General Approach: Cognition



- **Conceptualization and Place Classification** : Incremental process where objects perceived are grouped into clusters with each of them being inferred as an instance of a concept. These concepts are used to classify the place.

- Two processes: Clustering & Conceptualization (akin to segmentation and labeling)

# Clustering

- Based on distance and semantic models learnt.
- Distance
  - nearest neighbor ; distance to center of cluster
- Closest semantic group
  - MAP estimate of the best case concept suggested by the concept models learnt, for each perceived object
  - Uses models learnt (  $M_2$  ) as the best group implied by the occurrence of the single object is required.
  - $M_2$  modeled the likelihood of the occurrence of an object in an instance of a concept.

## clustering contd.

- Behavior (in same order)
  - Choose the nearest cluster that has the same concept as the best case concept
  - Choose the nearest cluster that is conceptually dissimilar but “acceptably likely” with respect to the best case concept.
  - Create a new cluster of type suggested by best case concept.

# Approach (M4): Details

- Naïve Bayes framework

$$P(c, X_{1...n1}, R_{1...n2}) = P(c) * \prod_{i=1}^{n1} P(X_i | c) * \prod_{j=1}^{n2} P(R_j | c)$$

- M4 = Object count (M3) + spatial relationships (distances)
- Object model:
  - M3 (object count) with a Gaussian uncertainty applied on the training data.
- Relationship model:
  - Gaussian mixture model learnt using EM algorithm.

# Object model

- Encodes the likelihood of the occurrence of a specific number of a particular object towards the formation of a particular concept.

$$P(o_i = m_i | c_i) = \frac{N_{o_i = m_i} + \delta}{N_{\text{exemplars}} + (2\delta)}$$

- A fixed Gaussian uncertainty added to training input. Every training exemplar will affect  $o_i = m_i - 1$ ,  $m_i$ ,  $m_i + 1$ . Attempt to improve generalization capability by handling conceptually adjacent cases.

# Relationship Model

- Choose relationships to be modeled
  - Based on number of occurrences
    - Both for significance and for practicality of GMM modeling
- EM based GMM modeling of relationship
  - Motivation for GMM – we want to capture characteristic relationships (values) between objects.
  - Use uncertain data (ie. relationship values with associated covariance matrices encoded in the representation).
  - BIC based model selection
- Check modeling for selection of relationship
  - Parametric bootstrapping in conjunction with a student's t-test (two tailed; 1 % level of significance) to check the modeling of the data.

# M4 : Bayesian Program

## Variables

$c, X_{1\dots n_1}, R_{1\dots n_2}$  where  $X_i \equiv o_i = m_i$   
i.e.  $X_i \equiv m_i$  occurrences of object  $o_i$  and  $c \equiv$  the concept  
and  $R_i$  is a relationship occurring between two objects

## Decomposition

$$P(c, X_{1\dots n_1}, R_{1\dots n_2}) = P(c) * \prod_{i=1}^{n_1} P(X_i | c) * \prod_{j=1}^{n_2} P(R_j | c)$$

## Parametric Forms

$$P(c) \rightarrow \begin{cases} P([c = 0]) = \frac{n_f + \delta}{n_f + n_t + 2\delta} \\ P([c = 1]) = \frac{n_t + \delta}{n_f + n_t + 2\delta} \end{cases}$$

$$P(X_i | c) \rightarrow \begin{cases} P([X_i = 0] | [c = 0]) = 1 - \frac{n_{fi} + \delta}{n_f + 2\delta} \\ P([X_i = 0] | [c = 1]) = 1 - \frac{n_{ti} + \delta}{n_t + 2\delta} \\ P([X_i = 1] | [c = 0]) = \frac{n_{fi} + \delta}{n_f + 2\delta} \\ P([X_i = 1] | [c = 1]) = \frac{n_{ti} + \delta}{n_t + 2\delta} \end{cases}$$

$$P(R_j | c) \rightarrow \begin{cases} P(R_j | c = 0) = \text{Uniform}(R_j) \\ \text{is a uniform distribution} \\ P(R_j | c = 1) = \text{GMM}(R_j) \\ \text{GMM}(x) = N(x; \pi_j, \mu_j, \Sigma_j) \\ \text{is the learnt Gaussian mixture model} \end{cases}$$

Identification  $\rightarrow$  Parameters learned during the training process

Question  $\rightarrow P(c | X_{1\dots n_1}, R_{1\dots n_2})$

# Experiments

- Physically measured dataset of 11 offices and 8 kitchens (19 places, 991 objects, 77 object types, 13 concepts types).
- Clustering outcomes
  - Object in correct cluster (~69 %)
  - Object fused with another cluster (~30 %)
  - Object in singleton cluster (~ 1%)
- Conceptualization outcomes
  - Object correctly conceptualized
  - Object in cluster that has not been classified
  - Free object in training assigned a label in testing
  - Object incorrectly conceptualized.

# Evaluation of Conceptualization

M3(obj count) ; Place Classification = 100 % • Nocc = 5, 10, 20, 30 tested.

Outcome	Cases	% (of classified)	% (of total)
Incorrect	175	18.5381	17.6589
Not classified	47	-	4.7427
Free Object	9	0.9534	0.9082
Correct	760	80.5085	76.6902

- Higher Nocc => fewer relationships => better performance
- M4 will reduce to M3 in limiting case (no relations).

M4(Nocc= 20) Place Classification = 100 %

Outcome	Cases	% (of classified)	% (of total)
Incorrect	126	14.4495	12.7144
Not classified	119	-	12.0081
Free Object	9	1.0321	0.9082
Correct	737	84.5183	74.3693

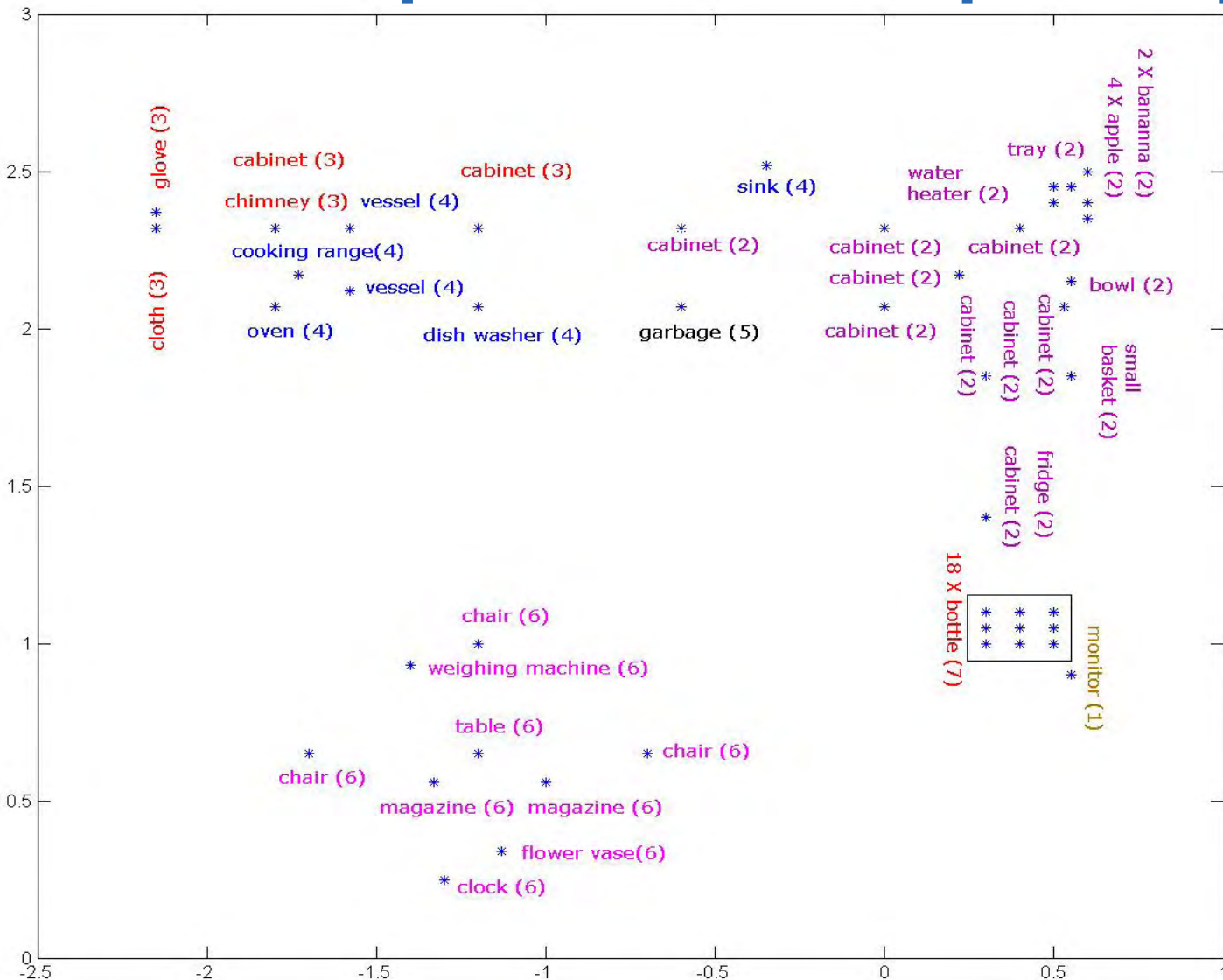
- Overall => reduction in number of incorrect cases (at the expense of more unclassified cases) and an increase in the classification accuracy (correctness amongst classified cases). for a similar number of correctly classified cases.

M4(Nocc= 30) Place Classification = 100 %

Outcome	Cases	% (of classified)	% (of total)
Incorrect	139	15.4102	14.0262
Not classified	89	-	8.9808
Free Object	9	0.9978	0.9082
Correct	754	83.5920	76.0848

- Inability to classify better than misclassification => improvement

# Experiments: Sample Output



Object occurrences only Cluster number - Output	
1	storage space
2	storage space
3	storage space
4	cooking space
5	garbage space
6	dining space
7	storage space (*)
Place Class = Kitchen	

Object + Relationships Cluster number - Output	
1	storage space
2	storage space
3	storage space
4	cooking space
5	garbage space
6	dining space
7	bottle group
Place Class = Kitchen	

Note on concepts

(\*) Incorrect concepts

(-) Unclassified (insufficient evidence)

# Further experiments on M3 and M4 (NOCC=20) K(=8)FSCV & LK(=2)OCV

KFSCV M3 accuracy - Expected Outcome

Outcome	Cases	% of Classified	% of Overall
Incorrect	327	33.7810	32.9970
Not classified	23	-	2.3209
Free Object	6	0.6198	0.6054
Correct	635	65.5992	64.0767

LKOCV M3 accuracy - Expected Outcome (across 88 tests)

Outcome	Cases	% of Classified	% of Overall
Incorrect	3487	37.1788	36.2361
Not classified	244	-	2.5356
Free Object	57	0.6077	0.5923
Correct	5835	62.2135	60.6360

KFSCV M4 accuracy - Expected Outcome

Outcome	Cases	% of Classified	% of Overall
Incorrect	248	27.7405	25.0252
Not classified	97	-	9.7881
Free Object	9	1.0067	0.9082
Correct	637	71.2528	64.2785

LKOCV M4 accuracy - Expected Outcome (across 88 tests)

Outcome	Cases	% of Classified	% of Overall
Incorrect	2588	30.0441	26.8939
Not classified	1009	-	10.4853
Free Object	72	0.8358	0.7482
Correct	5954	69.1200	61.8726

# Incorporation of Directional Relationships

Feature	Nocc	Incorrect	Correct	Not Classified	Places	Remarks
-	-	17.7	76.7	4.7	100	Object Count only
d	10	12.3	68.1	18.7	100	distance
	20	14.4	74.4	12	100	
	30	14	76.1	9	100	
txy	10	17.3	70	11.8	100	theta(xy)
	20	17.4	70.2	11.5	100	
	25	17.4	71	10.7	100	
	30	18.2	71.8	9.1	100	
tyz	10	19.5	58.5	21.4	94.7	theta(yz)
	20	18.6	71.4	9.4	100	
	25	20.4	72.6	6.2	100	
	30	18.3	75.4	5.4	100	
tzx	20	18.2	66.2	15	100	theta(zx)

Comparable (not yet better) results obtained – specially for txy and tyz.

# Outcome: Kitchen

## (movie of experiment on robot platform)



- Movie demonstrating thesis
  - Probabilistic Object graph representation of space
  - Bayesian Conceptualization & Place Classification
- Movie shows perception, representation / mapping, cognition.
  - inference using a-priori learnt models from a dataset

# Conclusions of C&C approach

- Bayesian approach conceptualization and place classification.
  - Based on Naïve Bayes framework; an object model that models the likelihood of the occurrence of a specific number of instances of an object in that of a concept; and a relationship model based on Gaussian mixture models.
  - Incorporating relationships improves performance of the approach.
    - Major effect : Reduce incorrect outcomes by not classifying them
    - Minor effect : Reduce incorrect outcomes by correctly classifying them
  - Grounded in underlying representation
  - Not ontology specific.
  - Generative and thus can enable other kinds of reasoning as well.

# User Studies

- **Human centered experiment to validate and enhance representation.**
- Quick overview
  - 52 people taken on a tour of lab premises
  - Questions on representation and description of objects, places, “zones” ; change of place, scene description etc.
  - categorize responses and infer ideas for sought representation
- Not a new theory of the mind / Cognitive Psychology
  - It does not answer the question of how the brain represents information.
  - Strictly an AI / robotics / engineering work.
- Results broadly validated approach.
- **Ref: Vasudevan@ICRA07-SIR (journal submission in review)**

# Some Results

- Cognitive = “human compatible”
- Object based representation of space is a viable solution.
- Objects are grouped into groups or concepts - semantic / functional abstractions in space.
  - similarities in purpose, functionality and also by the relative spatial arrangements of objects.
- Places can be understood as spatial abstractions
  - typically formed by bounding elements such as walls and doors.

## user study results contd.

- Change of Place
  - occurrence of boundary elements such as doors or walls and
  - a significant change in the kinds of objects and relative spatial arrangements of objects.
- Spatial abstractions contain semantic ones.
  - Indicated and subsequently taken as a design choice.
- A hierarchical representation of space is a good choice for the sought robot representation.
- and more...

# Why is this work **USEFUL** ?

- Context based disambiguation of objects
- Context sensitive (intelligent) path planning
- Intelligent HRI
  - Spatial / Social awareness
  - Can support / facilitate **INTELLIGENCE, INITIATIVE, PERSONALIZATION** and **HUMAN AUGMENTATION** – will be active themes of HRI research in the coming years.
- In addition to robot navigation.

# Contributions

## ■ Representation

- Hierarchical probabilistic concept-oriented representation of space.
  - Objects and spatial relationships between them.
  - Uncertainty is accounted for. Semantically enriched.
  - Human compatible (multi purpose) representation of space.

## ■ Cognition

- Generative Bayesian approach to conceptualization and place classification.
- Can map sensory information to increasingly abstract concepts towards creating a more semantic map for mobile robots.

## contributions contd.

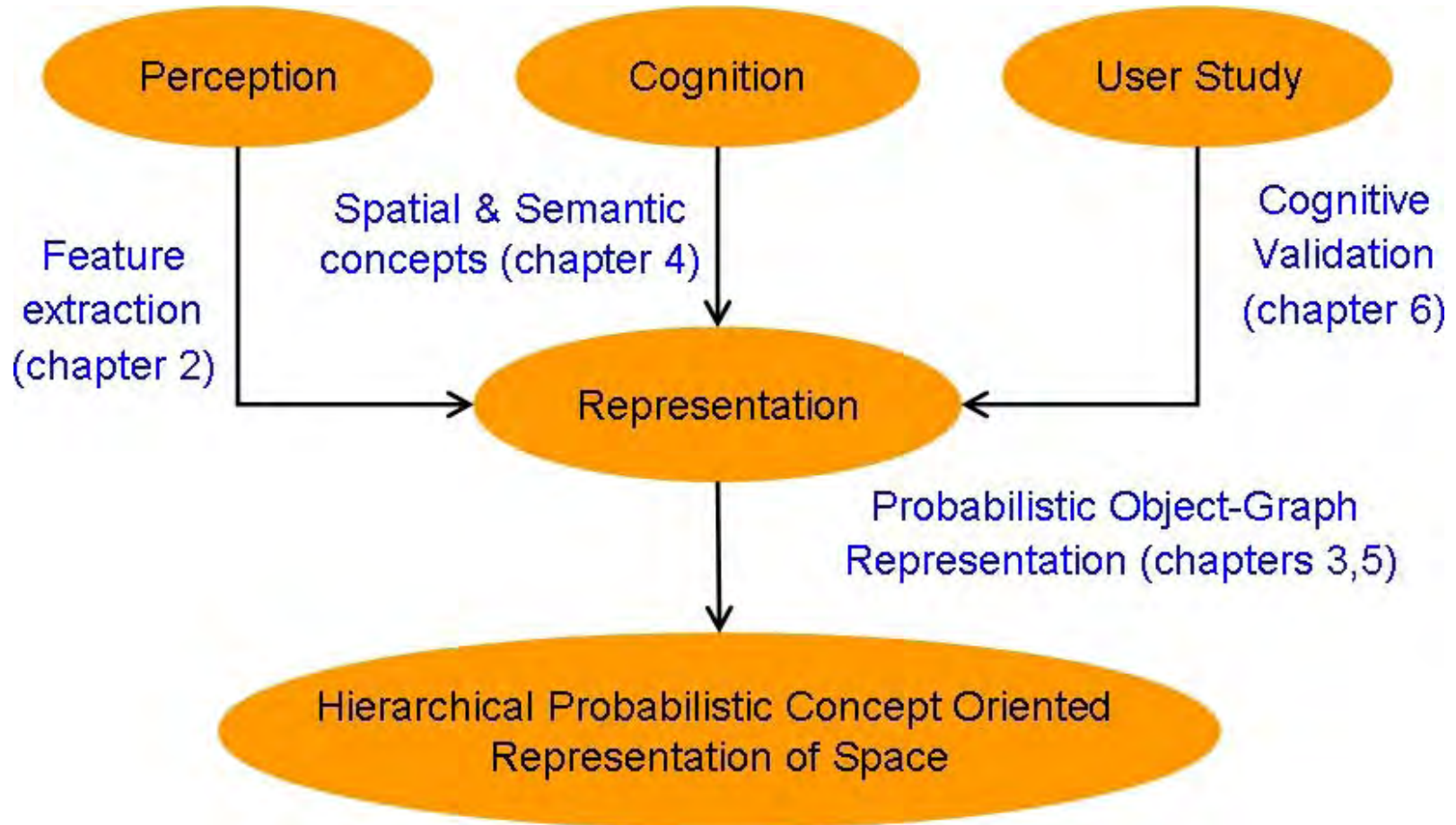
- User Studies – human centered validation of work
- Perception for real world implementation
- Application Outcome
  - Towards bringing robots into our homes.
- Scientific Outcomes
  - Increase in semantics in robot representations.
  - Increase in spatial awareness in robots.
  - Towards a new kind of SLAM (Semantic-Topological-Metric)
  - Towards bridging symbolic world (traditional AI) and Robotics

Abstraction level	Localization	Mapping
Semantic	Place Classification  Is this <b>AN</b> office ?	Concept Model
Topological	Place Recognition  Is this <b>MY</b> office ?	Topological Map
Metric	Metric Localization	Metric Map

The representation proposed in this thesis handles all 3 abstraction levels. The work can be easily extended into a new kind of Hierarchical (Semantic-Topo-Met) SLAM

Contribution: Extending SOA in mobile robotics to the Semantic level of abstraction

# Overview of Thesis



Shrihari Vasudevan, **Spatial Cognition for Mobile Robots: A Hierarchical Probabilistic Concept-Oriented Representation of Space**, Thesis Number 17612, March 2008, ETH Zurich, Switzerland.

## Related Publications (peer reviewed)

### ■ Book Chapters

- S. Vasudevan, S. Gächter, A. Harati and R. Siegwart (2007) **A hierarchical Concept-oriented Representation for Spatial Cognition in Mobile Robots**, In M. Lungarella, F. Iida, J. Bongard and R. Pfeifer (eds.), 50 Years of Artificial Intelligence, Springer Lecture Notes in Artificial Intelligence, Vol. 4850.

### ■ Journals

- S. Vasudevan, S. Gächter, C. Kobe and R. Siegwart (2008) **Cognitive Maps for Mobile Robots: Perspectives from a User Study**, Submitted.
- S. Vasudevan and R. Siegwart (2008) Bayesian Space Conceptualization and Place Classification for Semantic Maps in Mobile Robotics. **Accepted for publication in Robotics and Autonomous Systems.**
- S. Vasudevan, S. Gächter, V.T. Nguyen and R. Siegwart (2007) **Cognitive Maps for Mobile Robots - An object based approach**. Robotics and Autonomous Systems, Volume 55, Issue 5, From Sensors to Human Spatial Concepts, 31 May 2007, Pages 359-371.

### ■ Conferences & Workshops

- A. Ramisa, S. Vasudevan, D. Scaramuzza, R.L. de Mantaras and R. Siegwart (2008), **A tale of two object recognition methods for mobile robots**, Accepted for publication in the proceedings of the International Conference on Computer Vision Systems (ICVS), Santorini, Greece.

contd.

- S. Vasudevan and R. Siegwart (2007) **A Bayesian approach to Conceptualization and Place Classification: Incorporating Spatial Relationships (distances) to Infer Concepts**, in the proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'07) workshop "From Sensors to Human Spatial Concepts (FS2HSC), San Diego, USA
- S. Vasudevan, A. Harati and R. Siegwart (2007) **A Bayesian approach to Conceptualization and Place Classification: using the number of occurrences of objects to infer concepts**, in the proceedings of the 3rd European Conference on Mobile Robotics (ECMR) 2007, Freiburg, Germany.
- S. Vasudevan and R. Siegwart (2007) **A Bayesian Conceptualization of Space for mobile robots**. in the proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007), San Diego, USA. (more detailed technical report version below)
- S. Vasudevan, S. Gächter and R. Siegwart (2007) **Cognitive Spatial Representations for Mobile Robots – Perspectives from a user study**. In the proceedings of the IEEE International Conference on Robotics and Automation - Workshop: Semantic Information in Robotics (ICRA - SIR 2007), Rome, Italy.

contd.

- S. Vasudevan, S. Gachter, M. Berger and R. Siegwart (2006) **Cognitive Maps for Mobile Robots – An Object based Approach**. In the proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems - Workshop: From Sensors to Human Spatial Concepts (IROS - FS2HSC 2006), Beijing, China.
- S. Vasudevan, V.T. Nguyen and R. Siegwart (2006) **Towards a Cognitive Probabilistic Representation of Space for Mobile Robots**. In the Proceedings of the IEEE International Conference on Information Acquisition (ICIA) 2006, Shandong, China.
- A. Tapus, S. Vasudevan and R. Siegwart (2005) **Toward a Multilevel Cognitive Probabilistic Representation of Space**. In Proceedings of the International Conference on Human Vision and Electronic Imaging X, part of the IS&T/SPIE Symposium on Electronic Imaging 2005, 16-20 January 2005, CA, USA.
  
- **Links**
  - <http://www.asl.ethz.ch/research/asl/cogniron>
  - It has all above mentioned publications, movies & object recognition software.