A Generalized Query Framework for Geospatial Reasoning

Martin Michalowski
University of Southern California/
Information Sciences Institute
July 17th 2006

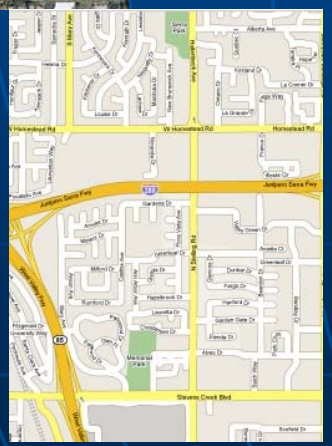
Introduction

- What is geospatial reasoning?
  - The fusion of traditional and non-traditional data sources

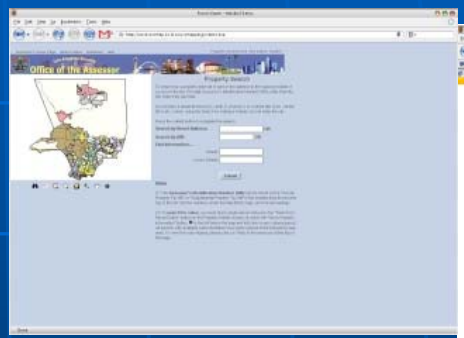
  AND
  - The ability to infer information from the integration of these sources

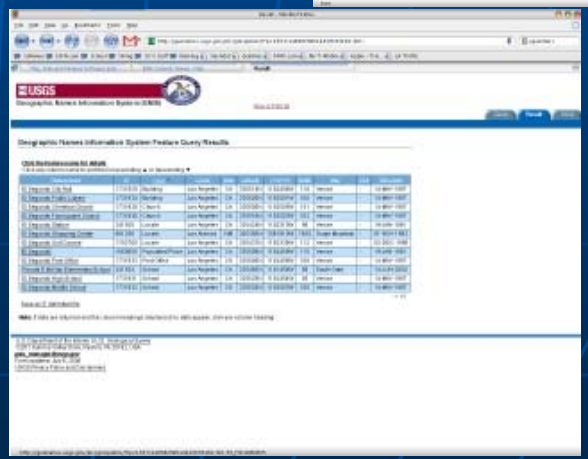
Introduction (cont.)

- Traditional data sources



- Non-traditional data sources





Previous Work

- Constraint Satisfaction Approach to Geospatial Reasoning (Michalowski and Knoblock AAAI-05)
  - Showed feasibility of planning techniques for geospatial data integration

  BUT

  - Abstracts away the problem of data gathering and model selection

Research Proposal

- Create a generalized query framework to answer queries about geospatial entities
  - Gathers appropriate data
  - Invokes appropriate tools (i.e. CSP solver)
  - Selects appropriate models to use
  - Provides complete and accurate answers to user queries

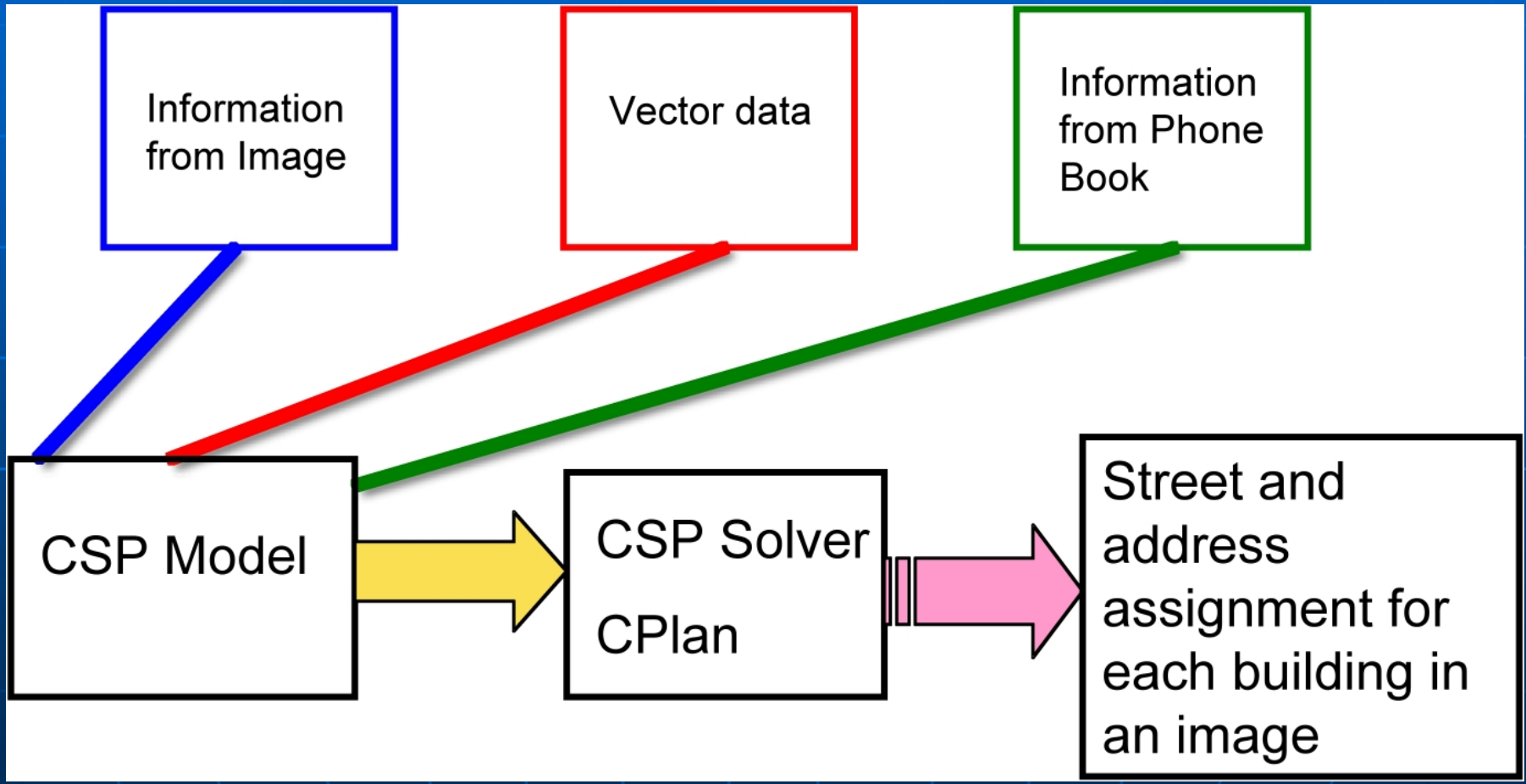
Motivating Example

El Segundo CA

Before

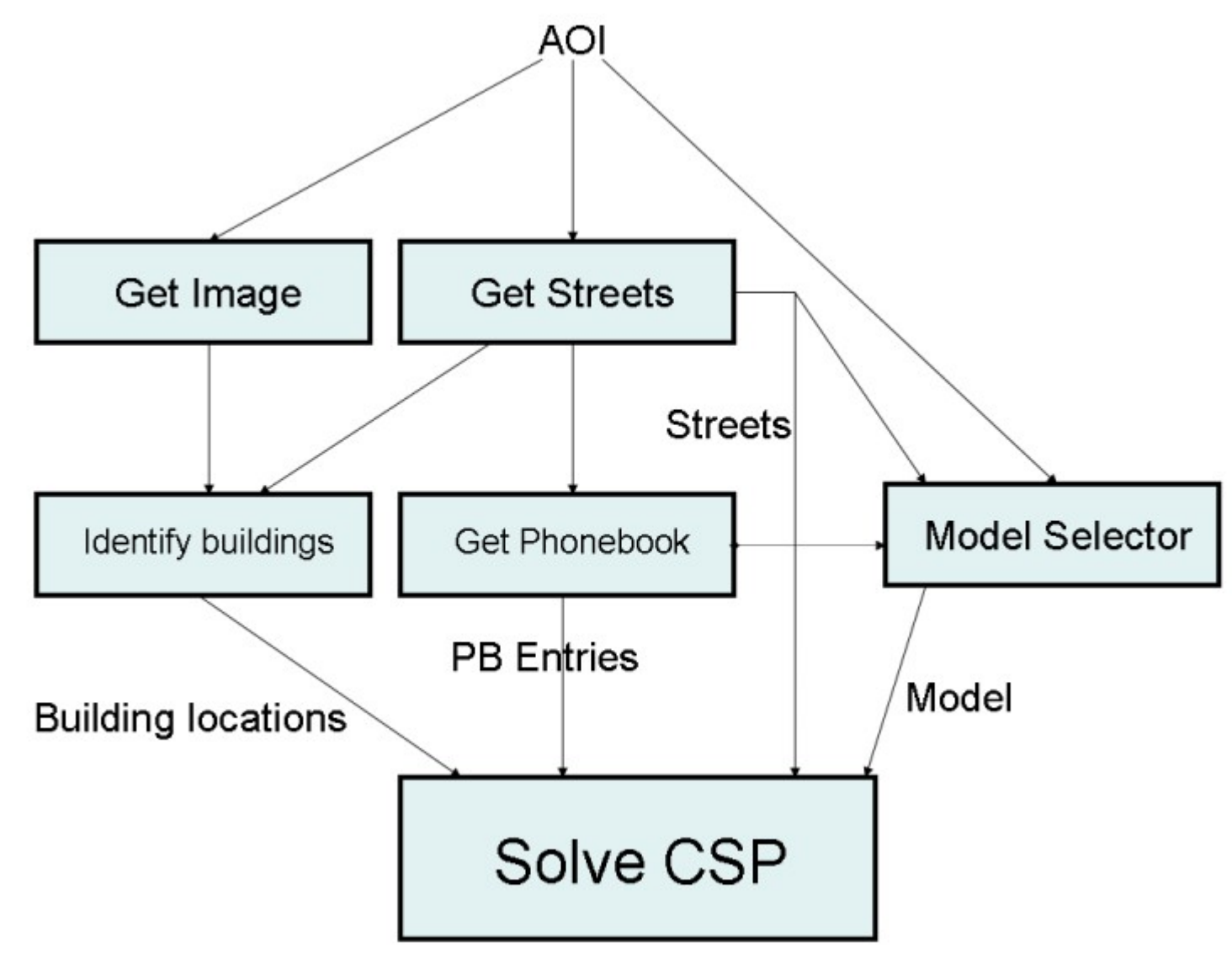
After

Integration Framework

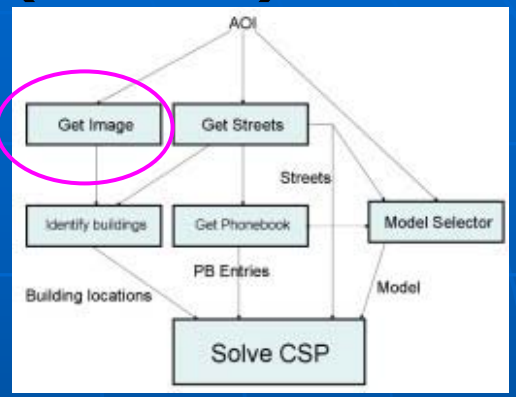


(Michalowski and Knoblock 2005)

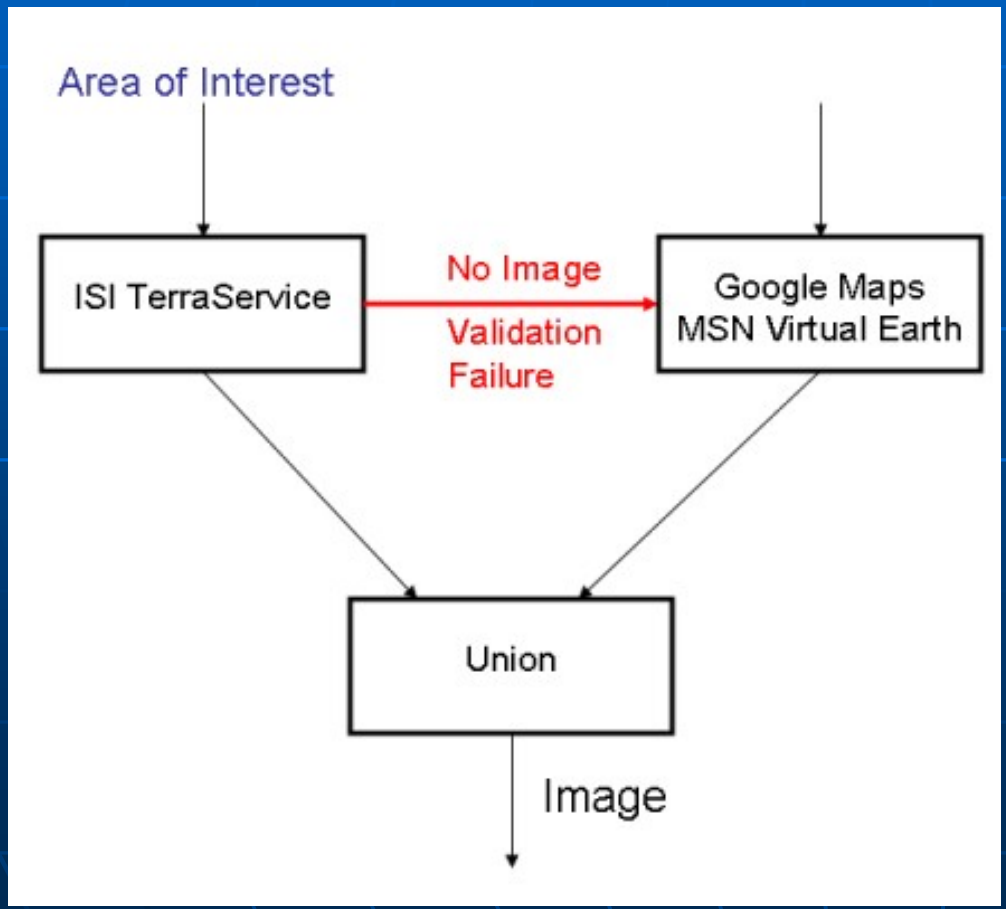
Composition Problem



Composition Problem (cont)



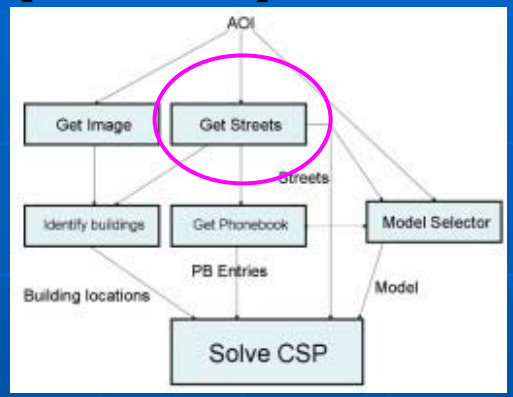
Get Image Component



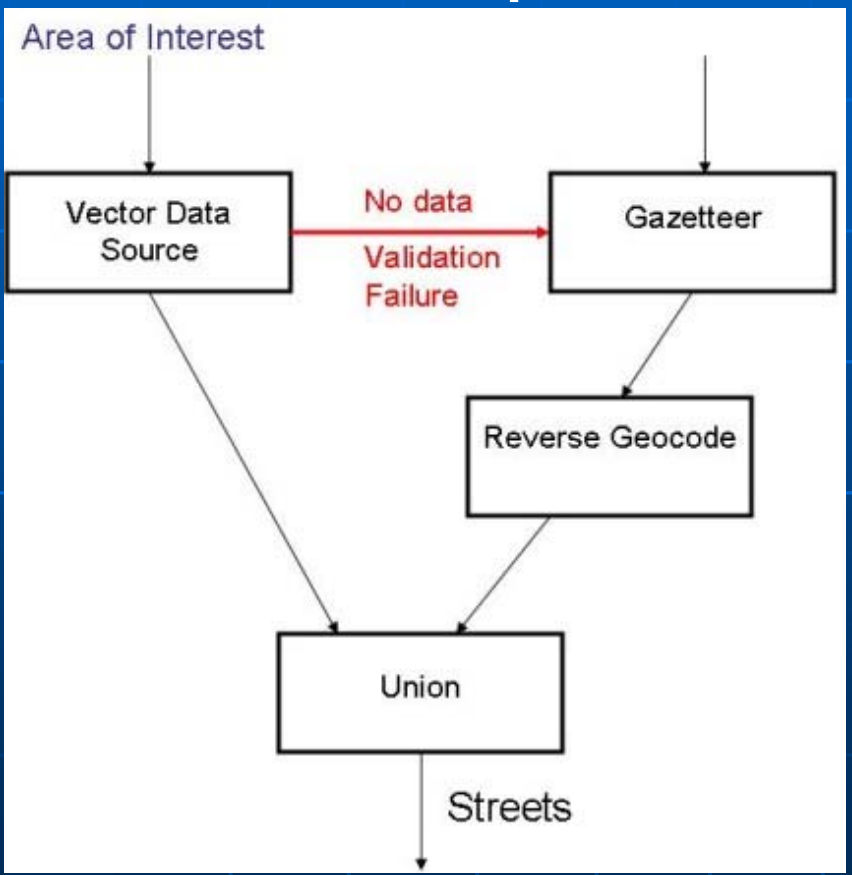
Retrieves the satellite imagery of El Segundo

Potential Issues:
- Image may not be available
- Different images may come from different sources -> which one to choose (operator utility)
- A source can produce multiple images

Composition Problem (cont)



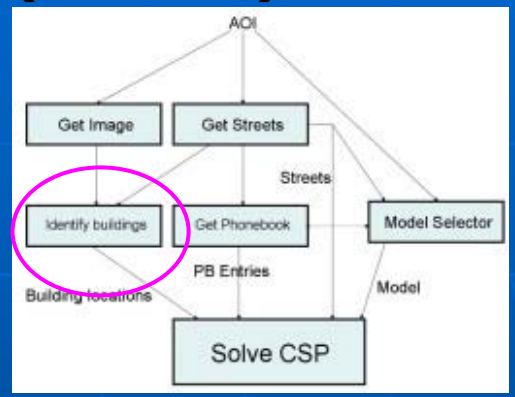
Get Streets Component



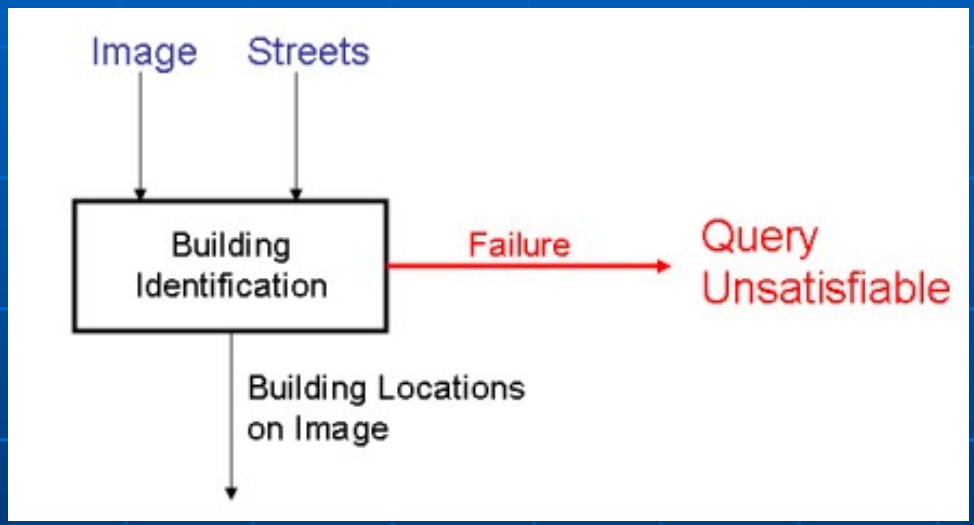
Gathers all of the street information about the AOI

Potential Issues:
- Not all streets may be in one source
- May need to query multiple sources to get a complete list of streets
- Multiple plans can produce street information -> which one to choose (plan utility)
- Street names can also be extracted from maps (Desai et. al 2005)

Composition Problem (cont)



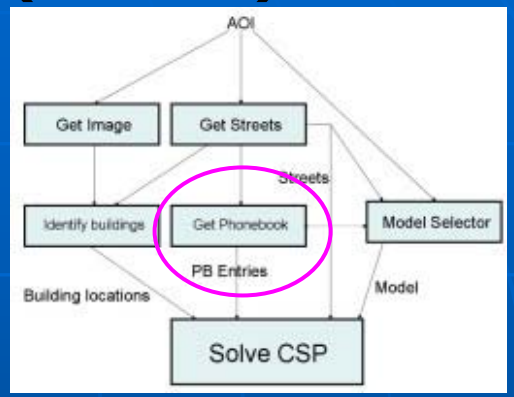
Identify Buildings Component



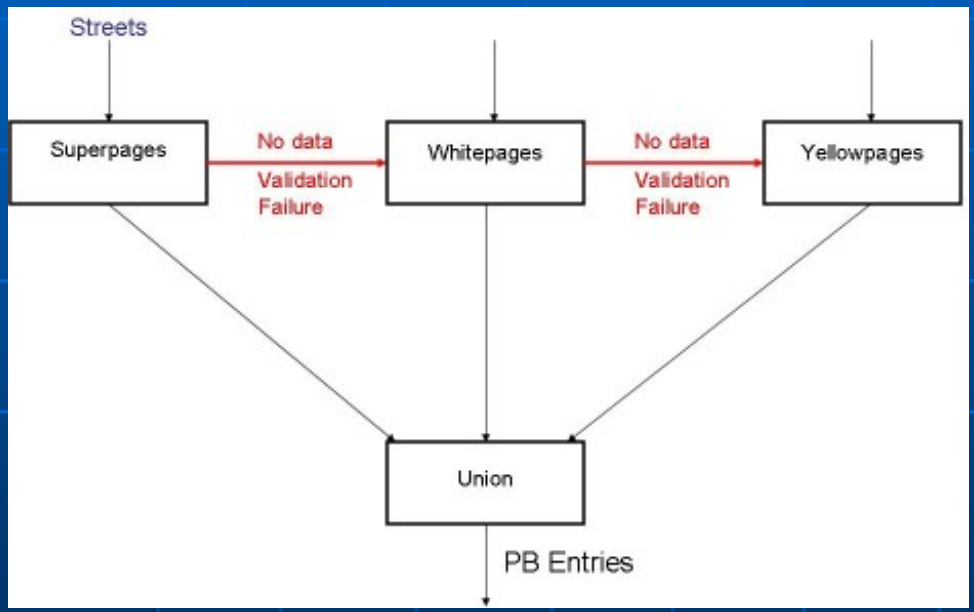
Uses the satellite imagery and street information to identify buildings in the imagery

Potential Issues:
- Correct tool needs to be invoked
- May not detect any buildings

Composition Problem (cont)



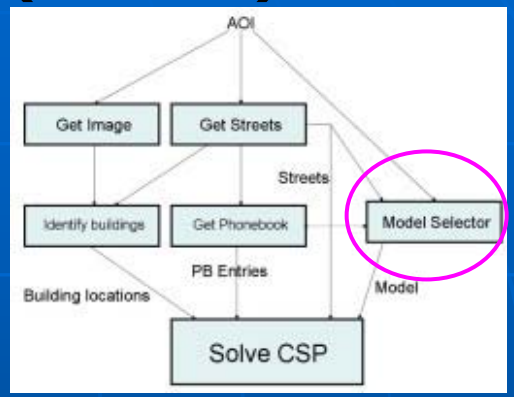
Get Phonebook Component



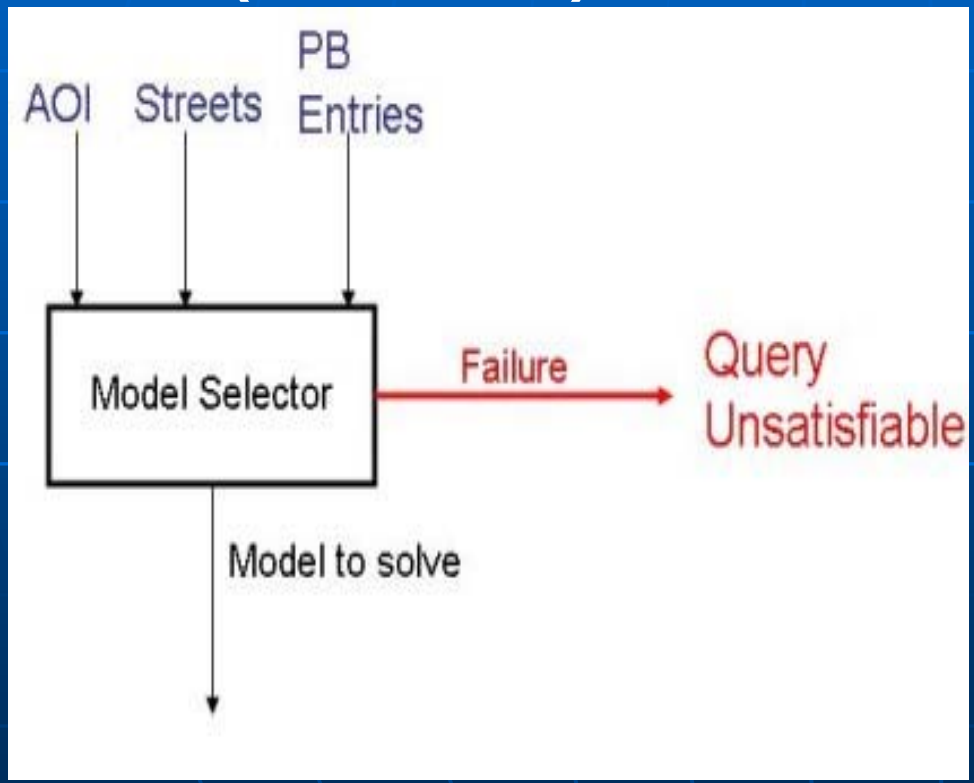
Uses the Superpages, Whitepages and/or Yellowpages to gather phone book information

Potential Issues:
- May need to query all sources to ensure completeness
- Some sources may be unavailable or fail for the given AOI
- Multiple plans can produce phone book information -> which one to choose (plan utility)

Composition Problem (cont)

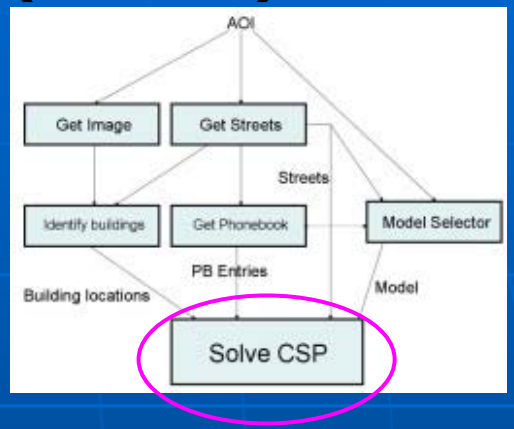


Model (Constraint) Selector Component

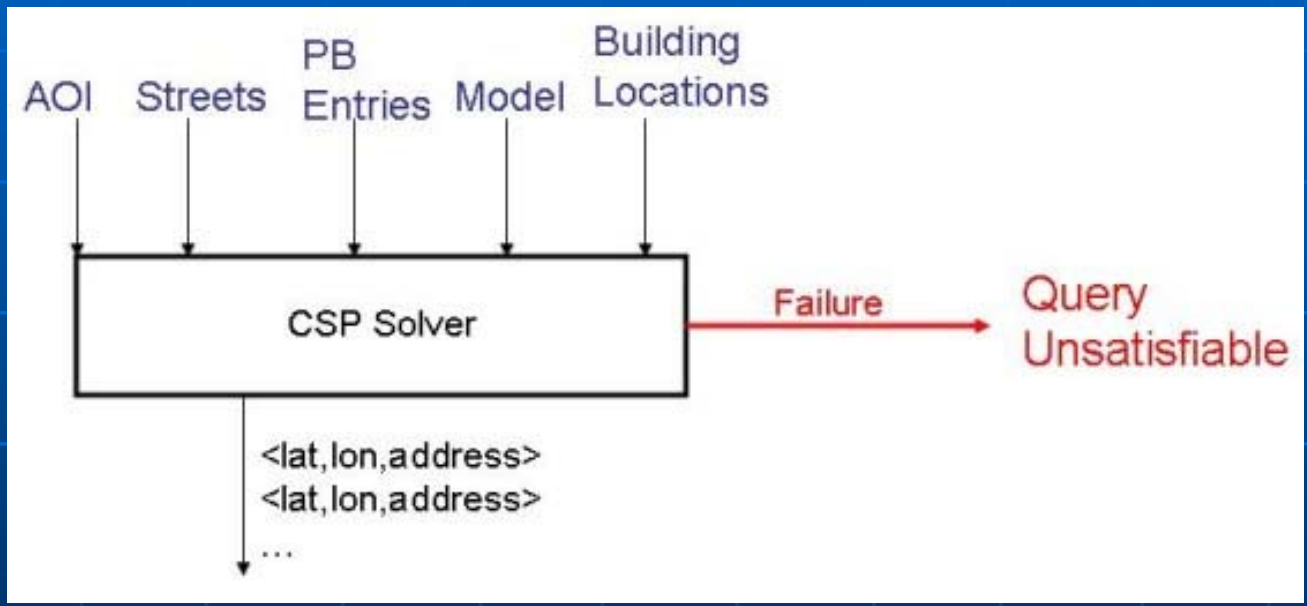


Combines information together to choose a CSP model
- Can refine the model by inferring information such as numbering schemes, address policies, etc.
- 100-blocking numbering applies to motivating example and can be inferred using:
- Landmark information (gazetteer)
- Hierarchy of numbering schemes
- Combined to infer that the block numbering scheme does apply

Composition Problem (cont)



Solve CSP Component



-Provides complete and accurate answers to user queries (if CSP is solvable)

-Can also augment solutions with additional information (information not requested in the query)

Outstanding Issues

- Different actions are available to the system
  - Need to model both the actions (operators) and the data produced
  - **Solution**: Hybrid approach of abstract and tuple level representations
- Need to deal with source/validation failures
  - Requires planning at the data AND operator level
  - **Solution**: Interleave planning and execution, re-planning only done with tuples that fail validation
- Data inconsistencies may arise across sources that provide the same information
  - Requires the use of utilities both at the plan and operator level
  - **Solution**: Conflict resolution strategies that combine multiple utility measures (both absolute and relative)

Outstanding Issues (cont.)

- CSP model can be selected based on available information
  - More detailed model leads to more accurate results
  - **Solution**: Try to infer certain constraints (i.e. numbering schemes) based on the available data
- Need to provide complete and accurate answers
  - Must ensure the user query is answered
  - **Solution**: Model selection, conflict resolution...

Related Work

- Planning for information gathering
  - IPEM (Ambros-Ingerson 88)
  - OCCAM (Kwok 94)
  - PUCCINI (Golden 98)
  - SAGE (Knoblock 95)
  - PKS (Petrick 04)
- Mediator Systems
  - SIMS (Arens 95)
  - InfoMaster (Dushka 97)
  - Prometheus (Thakkar 05)
- Conflict Resolution Strategies (Bleiholder 06)

Progress to Date

- Established feasibility of geospatial data integration using planning techniques
  - Developed a CSP framework to identify buildings in satellite imagery
- Study of related work in planning for information gathering
- Determining the correct representation
- Prototypical system implemented

Plan for Completion

- Fully understand real world problems
  - Inferring numbering schemes
  - Data retrieval and fusion
  - CSP model creation and selection
- Use this understanding to create a general framework that meets all requirements
  - Refine data and operator representation and planning and re-planning techniques

Contributions

- A general planning approach to data integration with incomplete and inconsistent data sources
- A robust planning approach to geospatial data integration
- Operators executed based on validation results
- Previous work on dealing with failures in planning is too general (complete failures)
  - My work would support failures at a finer grained level (tuple level)
- Validation introduces notion of “alternate” plans in the presence of incomplete sources

Open Questions

- Universal Planning vs. Dynamic Re-planning?
- Do other conflict resolution strategies make sense?
- Are answer accuracy and completeness good evaluation metrics?

Acknowledgements

- Craig Knoblock (Advisor)
- Mark Carman (Initial CSP Model)
- Berthe Choueiry and Ken Bayer (University of Nebraska – Scaling up and custom solver)
- Snehal Thakkar (All things geospatial)
- Officemates (All things non-research)
- Colleen van Lent (DC Mentor)

Thank you!