iCity Ontology CQ Summary

|  |  |
| --- | --- |
| **project 2.4 ILUTE** | |
| General Summary of Requirements:  -range of queries  -basic look-up, querying about specific individuals and activities related to parking; likely to gain insights / explore collected / simulated data sets  -arithmetic and geospatial queries to summarize data of interest over time periods, find trends and averages of certain occurrences  -one inference-based query looking at causation between activities (possibly different types) | |
| User-Specified CQs:  *As identified by project team members* | Abstracted CQs:  *i.e. taking away the parking-specific concepts, what do these queries “boil down” to? What is the fundamental nature of the questions being asked?* |
| **Basic Look-up** | |
| * TASHA Output review |  |
| **Look-up with Arithmetic/Geospatial Reasoning** | |
|  |  |

|  |  |
| --- | --- |
| **Parking TO DO (may capture some already via CUHK; to map TTS data if available?)** | |
| General Summary of Requirements:  -range of queries  -basic look-up, querying about specific individuals and activities related to parking; likely to gain insights / explore collected / simulated data sets  -arithmetic and geospatial queries to summarize data of interest over time periods, find trends and averages of certain occurrences  -one inference-based query looking at causation between activities (possibly different types) | |
| User-Specified CQs:  *As identified by project team members* | Abstracted CQs:  *i.e. taking away the parking-specific concepts, what do these queries “boil down” to? What is the fundamental nature of the questions being asked?* |
| **Basic Look-up** | |
| * How many curb-side / garage / surface parking spaces are available in zone X in time period Y? * How long is a person (from the TTS survey) parking, and how much are they paying? * Where are the parking ticket hot-spots by time of day? * Where are the locations where people get dropped off at a destination (a query on TTS data?) | * (count) What features are in zone X at time T? * What activities (trip, parking, parking payment) did person X participate in? What is the duration of activity X (parking)? What is the policy of parking spot P? * (count, sort) How many occurrences of activity X (parking tickets) occur in zone Y during some interval T? * Given a destination D of a trip T, where does the subactivity X of T (drop-off) occur? |
| **Look-up with Arithmetic/Geospatial Reasoning** | |
| * Where is the closest available parking spot to a user’s destination? * How much have parking ticket citations increased over time in downtown Toronto? * Are parking tickets more likely to be issued where there are fewer legal parking spots? * Where are the areas with the largest difference between parking supply and demand? | * What Feature X (parking spot) is located closest to the location Y of the feature Y that is the destination of Trip X? Where X has some status (i.e. “available”) at time T, where T is current or T is the estimated time of the end of trip X. * (count) How many occurrences of type X (parking ticket citation) are there in zone Y (downtown Toronto) during interval T? (compare) How many occurrences of type X are there in zone Y during some other interval T2 where T2 is before T and T2 and T have the same duration. * (count) How many occurrences of type X (parking ticket citation) occur in zone Y and how many legal parking spots are in zone Y? Pose this query for a number of zones with some criteria (e.g. same zone area, over same period of time). |
| **Inference** | |
| * Does illegal parking lead to lower traffic volumes on a roadway link where the illegal parking incident occurred? | * (Inference on causality) Does activity X lead to state Y? *Does* X cause Y? *Can* X *lead to* Y? |

|  |  |
| --- | --- |
| **Complete Streets – TO DO ?** | |
| General Summary of Requirements: specific lookup queries, summaries of observed/surveyed data. | |
| User-Specified CQs:  *As identified by project team members* | Abstracted CQs:  *i.e. taking away the parking-specific concepts, what do these queries “boil down” to? What is the fundamental nature of the questions being asked?* |
| **Basic Look-up** | |
| * How many transit vehicles will pass through my street? * How many conflicts will this cause with pedestrians, bicycles, cars? * What type of streetscape gets the highest user ratings (from our potential survey work)? * Where are the flows of car commuters when they walk from their parked car to their final destination? How many roads do they cross en route? * What are the width of each part of a streetscape (bike lanes, sidewalk, transit ROW, parking, av lane, HOV, etc.) | * How many activities (trips) made my mode X occur on an arc that is located on street segment Y (during time T)? * How many activities (conflicts with pedestrians, bicycles, cars) occur on a given arc? Occur with transit vehicles? * (order) (average/cumulative) How is a streetscape of type X rated? * Where is the trip segment (activity) with mode X and destination Y located? (what arcs does it occur on?) (more specifically, a trip segment that begins at a parking spot and is preceded by a trip segment with a vehicle mode) * What is dimension X (width/length/…) of the geometry of transportation complex Y (of type Z, e.g. bike lane, sidewalk)? |
| **Look-up with Arithmetic/Geospatial Reasoning** | |
| * What is the total bicycle km travelled on roads with bike lanes versus roads without bike lanes? | * *What is the length of all of the road segments (where the road segments are not accessed by bike only arcs) accessed by arcs, travelled by mode X (bicycle), (during time T)? and What is the length of all of the road segments (where the road segments are accessed by bike only arcs) accessed by arcs, travelled by mode X (bicycle), (during time T)?* |
| **Inference** | |
| * Is there a *connection between* traffic conflicts and traffic volume? | * (inference on *connection*) Is there a causal connection between a state where a traffic conflict has occurred and a state of reduced/increased traffic volume? or: summary of comparisons: What is the traffic volume on some road segment (arc, series of arcs, or collection of arcs in a particular area; TBD what the area of impact we are looking at is) at some time, T? and What is the traffic volume on the same road segment (or set of arcs) at some future time, after a conflict as occurred at location L? |

|  |  |
| --- | --- |
| **Transit Resilience** | |
| General Summary of Requirements:  Some specific data or statistical summary questions, but most of the queries are focused on causal relationships. For example, what types of things cause a particular type of state or type of activity or effect the way in which it occurs. | |
| User-Specified CQs:  *As identified by project team members* | Abstracted CQs:  *i.e. taking away the parking-specific concepts, what do these queries “boil down” to? What is the fundamental nature of the questions being asked?* |
| **Basic Look-up** | |
| * Where are these outdoor segments of the Toronto subway system? * How many buses were pulled from a specific bus route to serve in the shuttle service?  (to define: reroute activity) * ~~What are the types of streetcar and buses used in Toronto? And what are their average speed, acceleration and deceleration rates, number of seats, and crush (?) capacity?~~ | * What arcs (subway arcs) are underground (vs outdoor)? * How many transit vehicles participated in a specific class of activities? Data not available => use geospatial queries to attempt to detect rerouted buses * ~~Direct look-up of transit vehicle types and attributes~~   Data not available |
| **Look-up with Arithmetic/Geospatial Reasoning** | |
| * Which bus and streetcar routes are within 200 metre walking distance from subway stations?  (interpreted as: which routes have stops within 200m) | * What (routes containing) nodes have a location that is within 200 m of a given location(s) (subway station(s))? |

|  |  |
| --- | --- |
| **Transit Signal Prioritization** | |
| General Summary of Requirements:  Most queries are focused on determining how some activity can be used to achieve some state, with restrictions placed on either the type of activity or the type of state. | |
| User-Specified CQs:  *As identified by project team members* | Abstracted CQs:  *i.e. taking away the parking-specific concepts, what do these queries “boil down” to? What is the fundamental nature of the questions being asked?* |
| **Basic Look-up** | |
| * What are the traffic volumes, number of cyclists and pedestrians at TSP locations?   Q: do we have access to this data? | * What is the value of attribute X on the transportation network at location Y (at time T)? |

|  |  |
| --- | --- |
| **ATIS (project 1.2)** | |
| General Summary of Requirements:  Based on use cases identified bottleneck analysis, weather analysis, and for multimodal route selection, interactive road network information, and integration modules. | |
| User-Specified CQs:  *As identified by project team members* | Abstracted CQs:  *i.e. taking away the parking-specific concepts, what do these queries “boil down” to? What is the fundamental nature of the questions being asked?* |
| **Basic Look-up** | |
| * What are the attributes of some road segment, link, node, or some traffic event? * What is the current volume of some road link, and what is its capacity? | * What are the attributes of some road segment, link, node, or some traffic event? * What is the current volume of some road link, and what is its capacity? |

# Future work

|  |  |
| --- | --- |
| **Project 2.3 Transit Resilience - Basic Look-up (if data becomes available)** | |
| * Where are these outdoor segments of the Toronto subway system? * What are the types of streetcar and buses used in Toronto? And what are their average speed, acceleration and deceleration rates, number of seats,… | * What arcs (subway arcs) are underground (vs outdoor)? * Direct look-up of transit vehicle types and attributes |

|  |  |
| --- | --- |
| **Project 2.3 Transit Resilience - Inference** | |
| * Which streetcar routes suffer most from bunching, and where are the service bunching hot-spots by time of day? (given a classification of bunching) * What are the resources that comprise transit service resilience? (given some definition of transit service resilience) * What is the system average response and recovery time to major subway and streetcar service interruptions? * How do subway service interruptions affect the surface transit (i.e. bus and streetcar) systems in Toronto? * What are the local impacts (local impacts vs effects? “immediate effects”?) of these interruptions on the surface transit systems performance? * What are the effects of retracting buses from scheduled services to deploy as shuttle service in response to subway and streetcar service interruptions in Toronto? * What are the factors the affect transit user travel behaviour? * How can information provision and incentives influence transit riders travel behaviour and route selection (to divert them from system hotspots)? * What is transit service bunching? * What are the internal and external (internal vs external?) factors that affect the probability and time to first bunching incidents? * How do traffic incidents, signal timing and weather conditions impact service bunching? * How do scheduled headway and headway deviation at terminals impact service bunching? * What are the effects of outdoor track segments of the subway system on the frequency and duration of subway service interruptions? * What are the joint effects of the outdoor track segments and weather conditions (e.g., snow and rain precipitation)? * Given a current state of a streetcar line: will bunching occur? If so when? | * Given a classification/definition of bunching: infer trips segments in which some attributes (bunching) can be recognized, count at route stops over some intervals and sort by highest * What resources are required to enable some class of activities? * Classify activities according to some definition (e.g. recover, response) What is the duration of a specified class of activities? (average) * What states *caused* by some class of activities (service interruptions) are of a certain type (surface transit vehicles), or enable some activity in which a certain type of resource (surface transit vehicles) participates? * What are the attribute values (performance measures) of a resource (surface transit vehicle) that is caused to change by some activity (interruption)? * What are the effects of some class (and sequence, e.g. X then Y) of activities? * What is the class of states (factors) that causes any activity (behavior) to be performed by some class of actors (transit users)? * What (transit travel) activities are caused by a specific state type (information/knowledge)? * Transit service bunching to be defined as an activity; then we can pose queries to understand it – what are the attributes of this activity; what states does it cause, what states is it caused by? * Classification of factors wrt the activity: e.g. enabling states that are owned by some organization… What are the allowable changes in state for a certain activity to occur (i.e. under what conditions will a slightly different activity occur, or occur at a different time) * How does a change in some state impact some activity (if/how/when it occurs)? (this captures several of the CQs on how something impacts bunching, interruptions, etc.) * Predictive analysis: given some state X is it possible for an activity Y to occur, if so how? (summarize range of possibilities and occurrence times/durations) |

|  |  |
| --- | --- |
| **Project 2.3 TSP - Inference** | |
| * What are the strategies that could be used to minimize signal delays for buses and streetcars in Toronto? * Where are the locations that provide the maximum benefits of employing a transit signal priority (TSP) system in Toronto? * How can TSP be used to improve streetcar and bus systems speed and reliability? * What are the impacts of employing TSP on on non-transit modes users, i.e., auto users, cyclists and pedestrians? | * What activities cause a particular state (reduced delays)? What set of activities causes consistent sets of a state with a particular attribute (reduced delays)? * What are the locations of activities in a certain class of activities (TSP implementation) that cause the most/least of some attribute (e.g. delay time)? Or cause the highest/lowest number of states/activities that are classified as “good”/”bad”? * How do specific attributes of some class of activities impact their effect (or degree of) impact on some class of states? * What are the effects of some class of activities on a particular class of things (transportation system users)? |

|  |  |
| --- | --- |
| **Project 1.2 Inference CQs** | |
| * Is some set of observations (combination of states) usual or unusual? * What links are bottlenecks in the transportation system? (recurring vs non-recurring?) * How does some weather impact transportation system performance (e.g. volumes and speeds)? \*\*\*we didn’t have micro-weather information at the time, but we may have it via geoTAB. | * Based on some definition of usual and unusual states, classify some state in the transportation system as usual or unusual. * Based on some definition of a bottleneck: classify some link or node in the transportation system as a bottleneck. Are the manifestations of the time varying entity repeatedly classified as bottlenecks, or only exceptionally? * How does some type of weather (classify weather states) or weather event cause changes to the state of the transportation system (specifically wrt performance attributes). |

**Motivating Question Types**

* Abstraction
  + Event recognition: what events can we recognize from the model (or raw data) behavior? (at the micro, meso, or macro level)
  + Geographic abstraction
  + Demographic abstraction
* Cause & Effect
  + Explanation of simulation results
  + Solution suggestion (*computer-aided innovation)*:What can we do to achieve output X from the simulation model? What can we do to meet requirements 1,2,3 in the design?
  + Predictive: if we change X, what happens to the output?
  + Model design: what do we need to answer question X?
* Topology
  + Queries based on connection, containment, parthood, and position knowledge
    - Data integration
    - Reasoning (traveler DSS): what factors impact travel from A to B? reasoning about the relationship between travel and connections: Road and lane closures (on Arcs), transit signal policies (at Nodes)
* Common Sense
  + Validation (models, analysis, collected data)
* Integration/Fusion:
  + What CQs are embedded in these problems? E.g. (re-)classification, translation between different geo-location systems, different model representations, …

# Waterfront Toronto

Simulation: travel demand model

* What input data do I need to explore a particular question? (Where) is it available?
* Cause/effect: (what) does changing feature x change in the results of the demand model output? What events of interest must have happened, given the model output? What caused change Z?
* Parthood:
* Connection:

3D modelling, visualization: exploring design alternatives

* Is there some feature common to all alternatives? Is there some feature not present in (/not required by) any alternatives?
* Is this design consistent with my rules? With the rest of the road network?
* Is there a path for mode X from points Y to Z?

# Simulation

can we infer an impact on the results without re-running a simulation?

* Goal-based suggestions (If we want to achieve X, what should the initial state be?)
* Question-based simulation design: I want to know the impact of adding a subway station on property values; what do I need?  
  OR, I want to know the impact of adding a subway station – what events can this cause?
* Results analysis (

# Data Collection

-IT-SoS

-Surveys

-Land use data

# Analysis...

-transit reliability

-signal management

-parking

-freight

-complete streets

## Data Management

The iCity team wants to keep better track of their data. To achieve this, they are creating a (set of) unified databases to store all of their data in a common location(s). Once the data is there, the key requirement will be to find (and understand) relevant datasets.

The ontology can support this in several ways:

1. Define metadata, provenance information so that this can be maintained uniformly and queried across the databases. “What data is available about parking in Zone-X?”
2. Support specific queries *into* the datasets, as opposed to about them. “Where are all of the parking spots in Zone-X?” “What are the different parking policies in Zone-X?”  
   This could simply involve the retrieval of a dataset that matches the required information, or it could involve the aggregation of several datasets that match, or (most interestingly) the integration of several datasets to create the required results.

Type 1 support requires the addition of metadata concepts to the ontology. There are existing vocabularies (possibly ontologies) to support this, e.g. W3C’s DCAT, Dublin Core, and so on. Additional domain-specific concepts will be dictated by the team (e.g. geospatial information, data privacy agreements and so on). Multiple versions of the same dataset must also be addressed.

Type 2 support requires consideration of the relationship between the real-world concepts as defined in the ontology, and the datasets. It is important that the user be able to understand where the answer to their query came from (provenance), i.e. which dataset(s) were used.