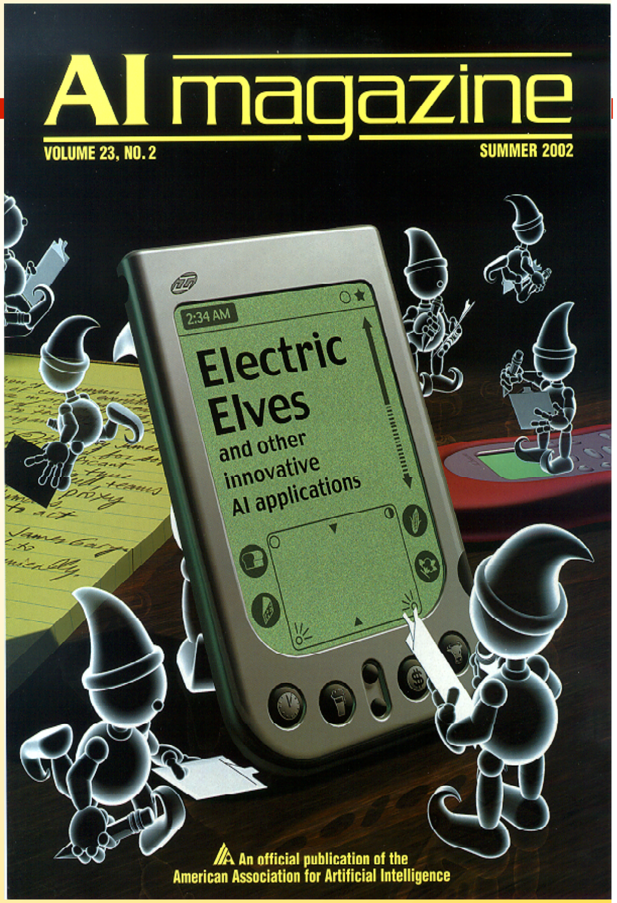
Beyond the Elves: Making Intelligent Agents Intelligent

Craig Knoblock

University of Southern California
Information Sciences Institute


AI magazine
VOLUME 23, NO. 2
SUMMER 2002
2:34 AM
Electric Elves
and other innovative AI applications
An official publication of the American Association for Artificial Intelligence

Contributors

- José Luis Ambite, USC
- Greg Barish, Fetch Technologies
- Hans Chalupsky, USC
- Yolanda Gil, USC
- Kristina Lerman, USC
- Steven Minton, Fetch Technologies
- Maria Muslea, USC
- Jean Oh, CMU
- David V. Pynadath, USC
- Thomas A. Russ, USC
- Paul Scerri, USC
- Milind Tambe, USC

Research Support

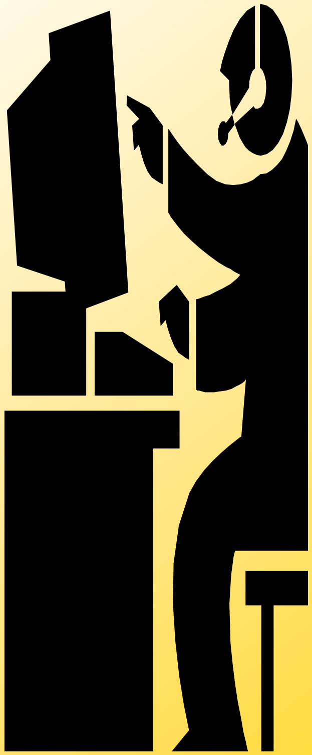
- Electric Elves
  - DARPA CoABS program
  - DARPA Active Templates program
- Research based on lessons learned
  - DARPA PAL Program
  - AFOSR

Outline

- The Electric Elves
  - Overview
- The Travel Elves
  - Constraint-based data integration
  - Efficient plan execution
- Beyond the Elves
  - Deployment to DARPA
  - Some lessons
    - Learning expectations
    - Finding alternative sources
    - Predicting prices
- Discussion

Agent-Facilitated Human Organizations

WWW
Ontology-Based Matchmakers
Agent Proxies For People
Information Agents



Goal

Agent-Facilitated Human Organizations

- Develop software agents that automate routine tasks within organization
  - E.g., obtaining info, monitoring, distributing information
- Enable software agents and humans to act coherently within the organization
  - Humans have agent proxies that assist in coordinating with other agents
- Coordination of tasks within the organization
  - Efficient use of resources (physical and human)
  - Monitoring progress of both individual and overall tasks
  - Suggest/execute corrective actions when goals appear endangered

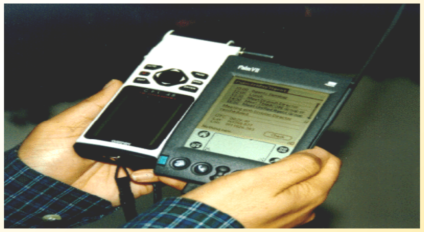
Example Tasks

Supporting our own research organization:

- Coordinating routine project activities
  - Eliminate email about meeting delays/postponements
  - Eliminate email about who will present at next meeting
  - Track visitor’s flights and arrival times
- Support for external meetings
  - Monitor all aspects of a traveler’s plans
    - Notify traveler of prices changes, schedule changes, flight delays
    - Send fax to hotel if traveler is delayed
  - Organize meetings for people with similar research interests in other organizations
  - Organize meeting at external site (e.g., a restaurant)
- Similar issues arise in many organizations: corporations, military, etc.

Agent Organization is Grounded in the Real World

- On-line information sources
  - Calendar and schedule information
  - Flight schedules, restaurants, etc.
- “Sensors” and “actuators”
  - GPS, “finger”, infrared, etc.
  - Postpone/cancel meetings, faxes to hotels and restaurants, etc.
- Interaction with people
  - Communicate through personal portable devices (PDAs, phones, etc.)



Summary

- Electric Elves: successful use of agent technology to support human organization
  - Interact with their environment:
    - Data: on-line information sources, calendars, GPS, etc.
    - Interaction with people: faxes, email, PDAs, phones, etc.
  - Support routine activities
    - Arrange meetings, lunch orders
    - Monitoring and suggesting/executing corrective actions
- Integrates powerful techniques in diverse areas within AI:
  - Teamwork, human proxies [Teamcore]
  - Ontologies & problem solving [Expect]
  - Information agents [Ariadne]
  - Knowledge representation & information retrieval [PowerLoom]

Outline

- The Electric Elves
  - Overview
- The Travel Elves
  - Constraint-based data integration
  - Efficient plan execution
- Beyond the Elves
  - Deployment to DARPA
  - Some lessons
    - Learning expectations
    - Finding alternative sources
    - Predicting prices
- Discussion

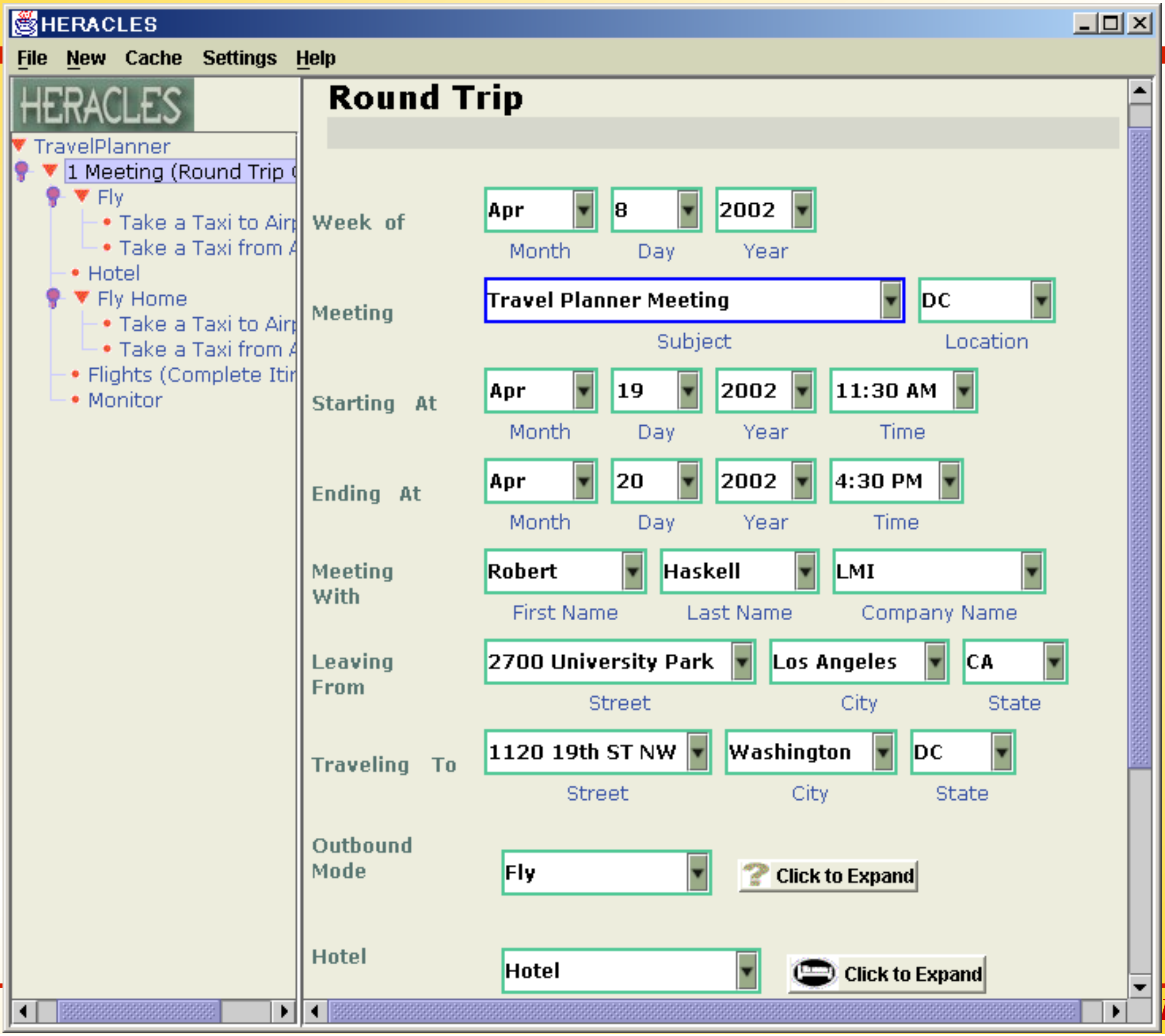
Interactive Trip Planning

- Commercial systems provide support to select flights, hotels and cars
  - Integrates the planning at the level of dates and locations
- There are many more factors involved in planning a trip
  - Which airports to fly into and out of
  - Whether to drive or take a taxi to the airport
  - How to get from the airport to the destination
  - Proximity of hotel to meeting
  - Etc…
- Ideally a system will
  - Provide all of the data required to make these decisions
  - Provide a way to consider the tradeoffs of the various choices

Heracles Constraint-based Planning

- Framework for building integrated applications
- Extract and integrate data for a given task
  - Live access to online sources using the wrappers
- Constraint-based system decides what sources to query and how to integrate the results
  - Tight integration of user choices

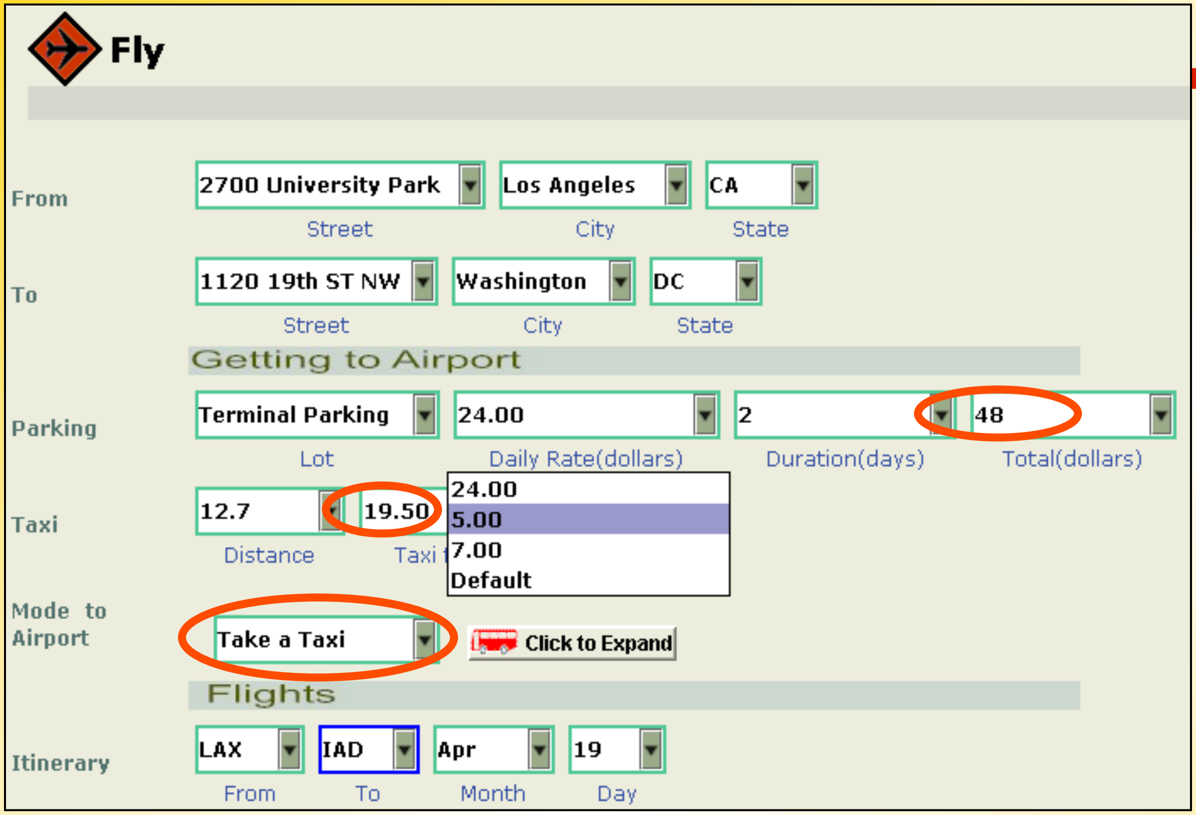
The Travel Assistant



Round Trip

Week of	Apr	8	2002	
	Month	Day	Year	
Meeting	Travel Planner Meeting		DC	
	Subject		Location	
Starting At	Apr	19	2002	11:30 AM
	Month	Day	Year	Time
Ending At	Apr	20	2002	4:30 PM
	Month	Day	Year	Time
Meeting With	Robert	Haskell	LMI	
	First Name	Last Name	Company Name	
Leaving From	2700 University Park	Los Angeles	CA	
	Street	City	State	
Traveling To	1120 19th ST NW	Washington	DC	
	Street	City	State	
Outbound Mode	Fly			
Hotel	Hotel			

Supports Informed Choices



Fly

From	2700 University Park	Los Angeles	CA	
	Street	City	State	
To	1120 19th ST NW	Washington	DC	
	Street	City	State	

Getting to Airport

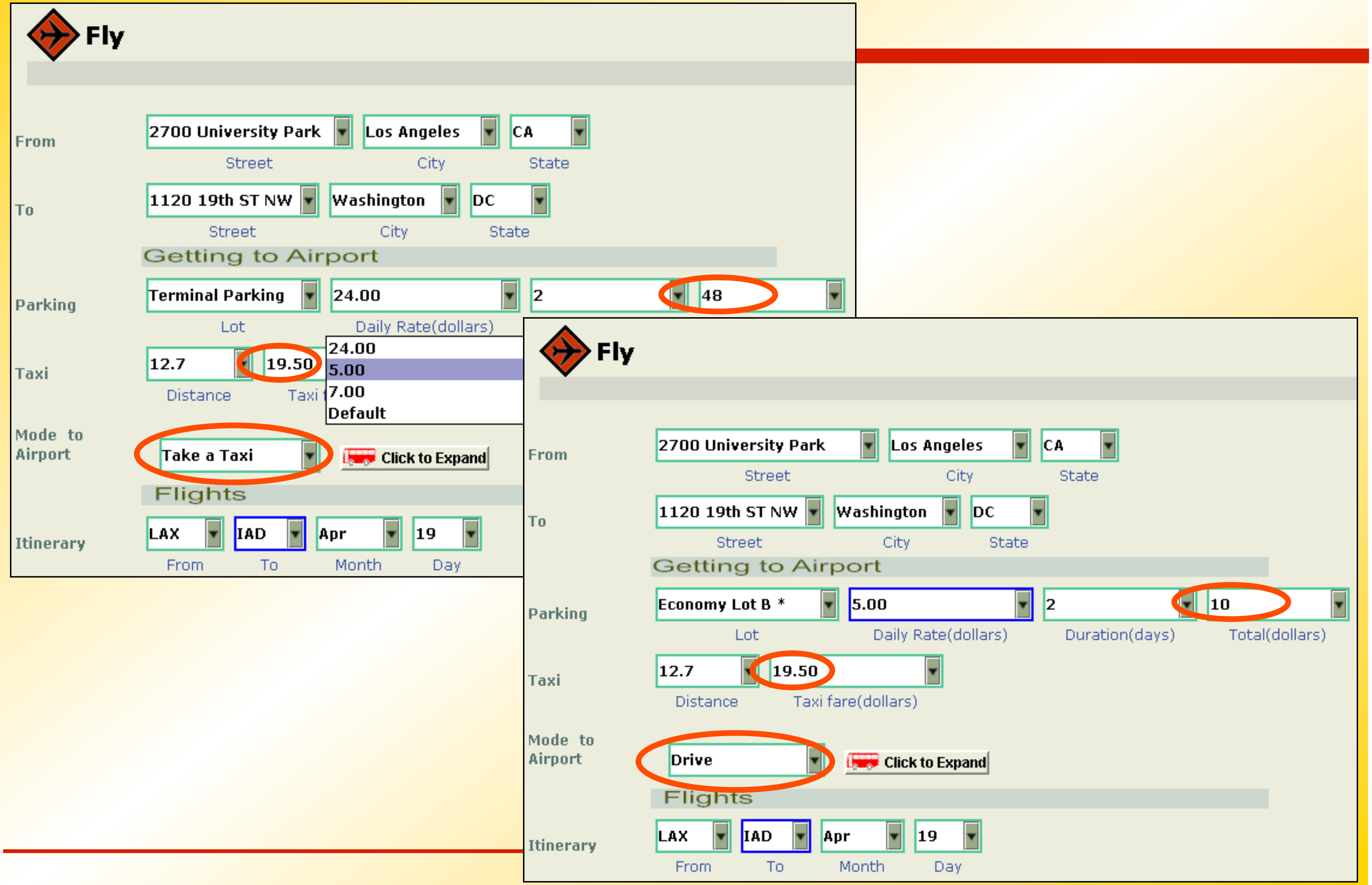
Parking	Terminal Parking	24.00	2	48
	Lot	Daily Rate(dollars)	Duration(days)	Total(dollars)
Taxi	12.7	19.50		
	Distance	Taxi f		
Mode to Airport	Take a Taxi			

Flights

Itinerary	LAX	IAD	Apr	19
	From	To	Month	Day

Supports Informed Choices

e-Elves



Fly

From: 2700 University Park (Street) | Los Angeles (City) | CA (State)

To: 1120 19th ST NW (Street) | Washington (City) | DC (State)

Getting to Airport

Parking: Terminal Parking (Lot) | 24.00 (Daily Rate(dollars)) | 2 | 48

24.00
5.00
7.00
Default

Taxi: 12.7 (Distance) | 19.50 (Taxi f

Mode to Airport: Take a Taxi

Flights

Itinerary: LAX (From) | IAD (To) | Apr (Month) | 19 (Day)

Fly

From: 2700 University Park (Street) | Los Angeles (City) | CA (State)

To: 1120 19th ST NW (Street) | Washington (City) | DC (State)

Getting to Airport

Parking: Economy Lot B * (Lot) | 5.00 (Daily Rate(dollars)) | 2 (Duration(days)) | 10 (Total(dollars))

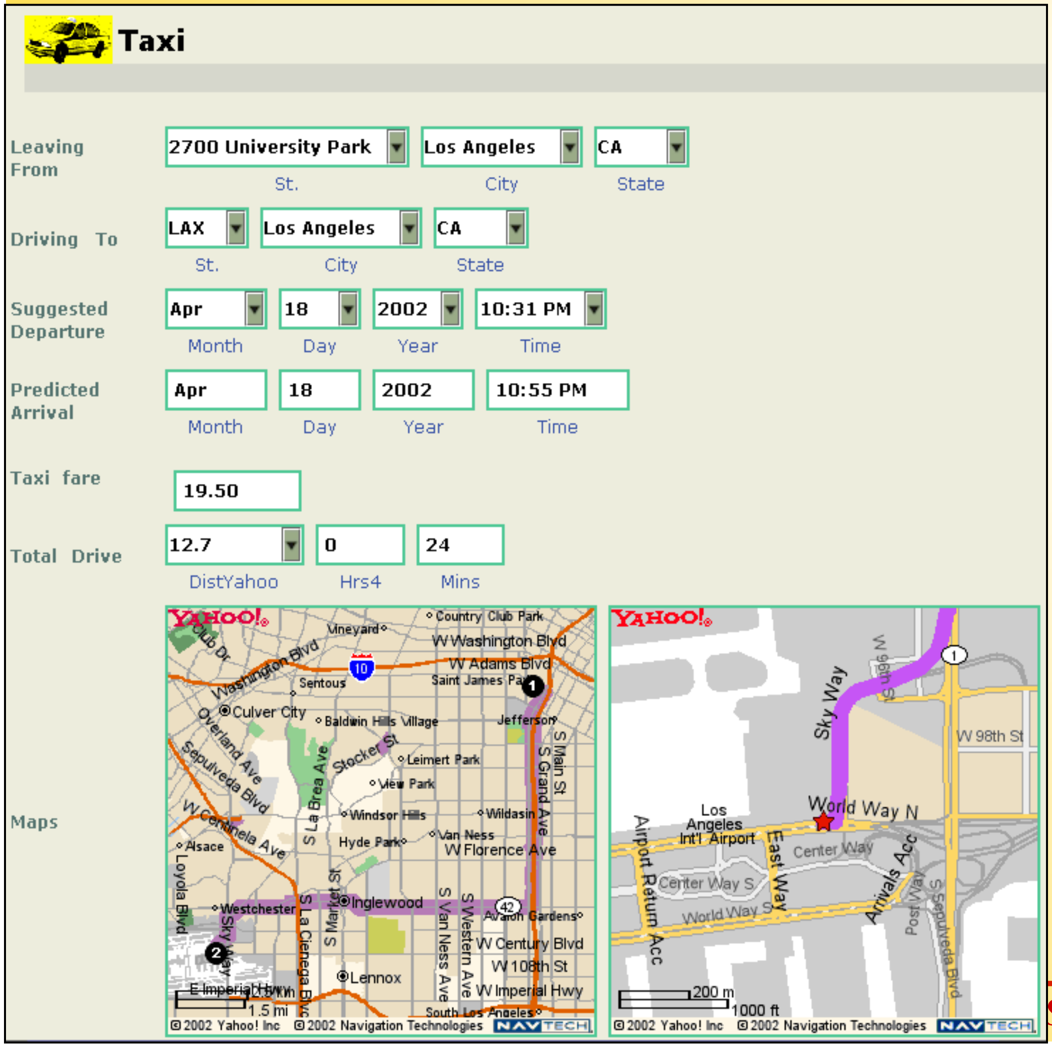
Taxi: 12.7 (Distance) | 19.50 (Taxi fare(dollars))

Mode to Airport: Drive

Flights

Itinerary: LAX (From) | IAD (To) | Apr (Month) | 19 (Day)

Changes Propagate Throughout



Taxi

Leaving From: 2700 University Park (St.) | Los Angeles (City) | CA (State)

Driving To: LAX (St.) | Los Angeles (City) | CA (State)

Suggested Departure: Apr (Month) | 18 (Day) | 2002 (Year) | 10:31 PM (Time)

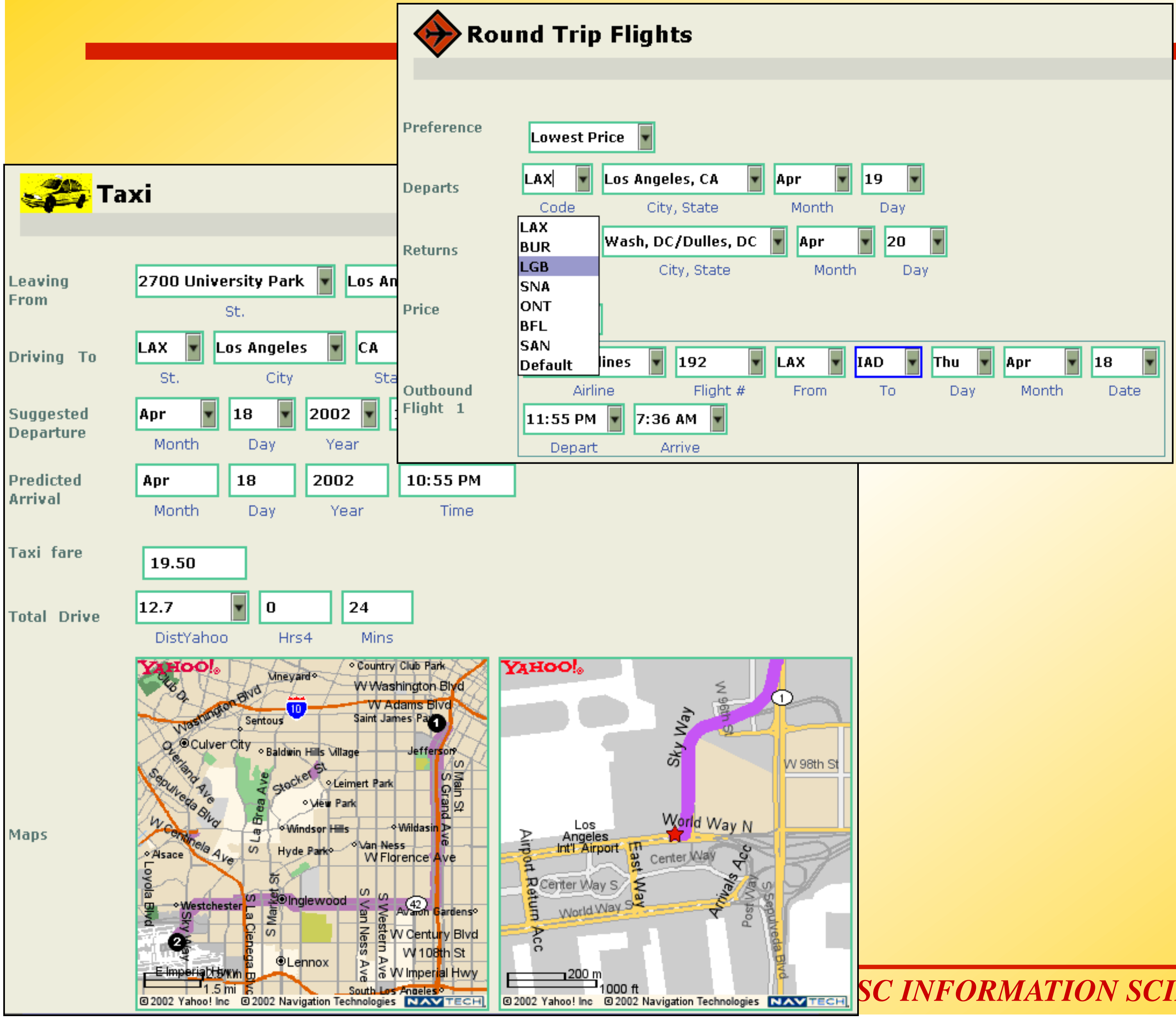
Predicted Arrival: Apr (Month) | 18 (Day) | 2002 (Year) | 10:55 PM (Time)

Taxi fare: 19.50

Total Drive: 12.7 (DistYahoo) | 0 (Hrs4) | 24 (Mins)

Maps

Changes Propagate Throughout



Round Trip Flights

Preference: Lowest Price

Departs: LAX | Los Angeles, CA | Apr | 19
Code | City, State | Month | Day

LAX
BUR
LGB
SNA
ONT
BFL
SAN
Default

Returns: Wash, DC/Dulles, DC | Apr | 20
City, State | Month | Day

Price

Outbound Flight 1: lines | 192 | LAX | IAD | Thu | Apr | 18
Airline | Flight # | From | To | Day | Month | Date
11:55 PM | 7:36 AM
Depart | Arrive

Taxi

Leaving From: 2700 University Park | Los An
St.

Driving To: LAX | Los Angeles | CA
St. | City | Sta

Suggested Departure: Apr | 18 | 2002
Month | Day | Year

Predicted Arrival: Apr | 18 | 2002 | 10:55 PM
Month | Day | Year | Time

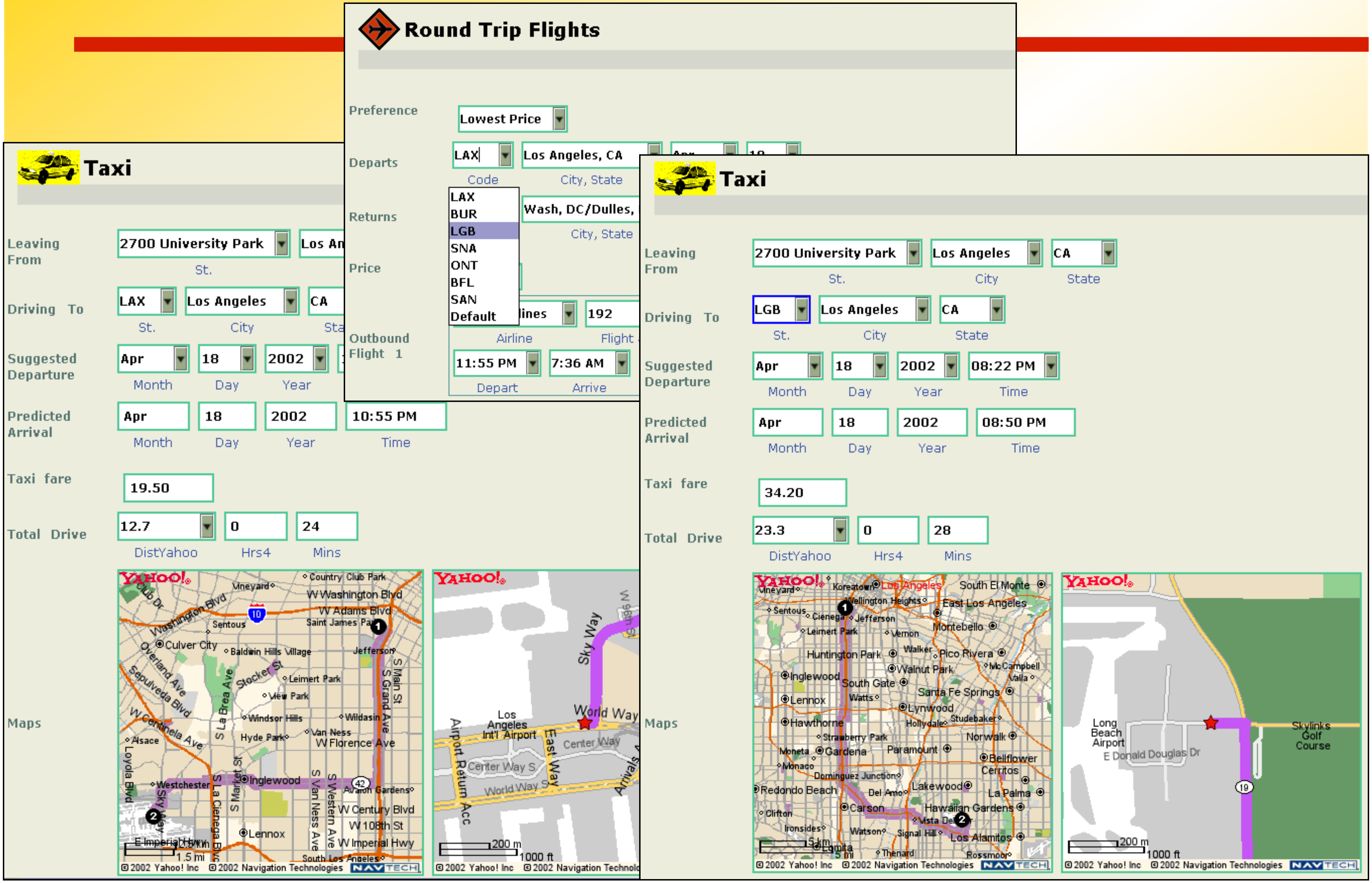
Taxi fare: 19.50

Total Drive: 12.7 | 0 | 24
DistYahoo | Hrs4 | Mins

Maps

e-Elves

Changes Propagate Throughout



Taxi

Leaving From: 2700 University Park | Los An
St.

Driving To: LAX | Los Angeles | CA
St. | City | Sta

Suggested Departure: Apr | 18 | 2002
Month | Day | Year

Predicted Arrival: Apr | 18 | 2002 | 10:55 PM
Month | Day | Year | Time

Taxi fare: 19.50

Total Drive: 12.7 | 0 | 24
DistYahoo | Hrs4 | Mins

Maps

Round Trip Flights

Preference: Lowest Price

Departs: LAX | Los Angeles, CA | Apr | 18
Code | City, State

LAX
BUR
LGB
SNA
ONT
BFL
SAN
Default

Returns: Wash, DC/Dulles,
City, State

Price

lines | 192
Airline | Flight

Outbound Flight 1: 11:55 PM | 7:36 AM
Depart | Arrive

Taxi

Leaving From: 2700 University Park | Los Angeles | CA
St. | City | State

Driving To: LGB | Los Angeles | CA
St. | City | State

Suggested Departure: Apr | 18 | 2002 | 08:22 PM
Month | Day | Year | Time

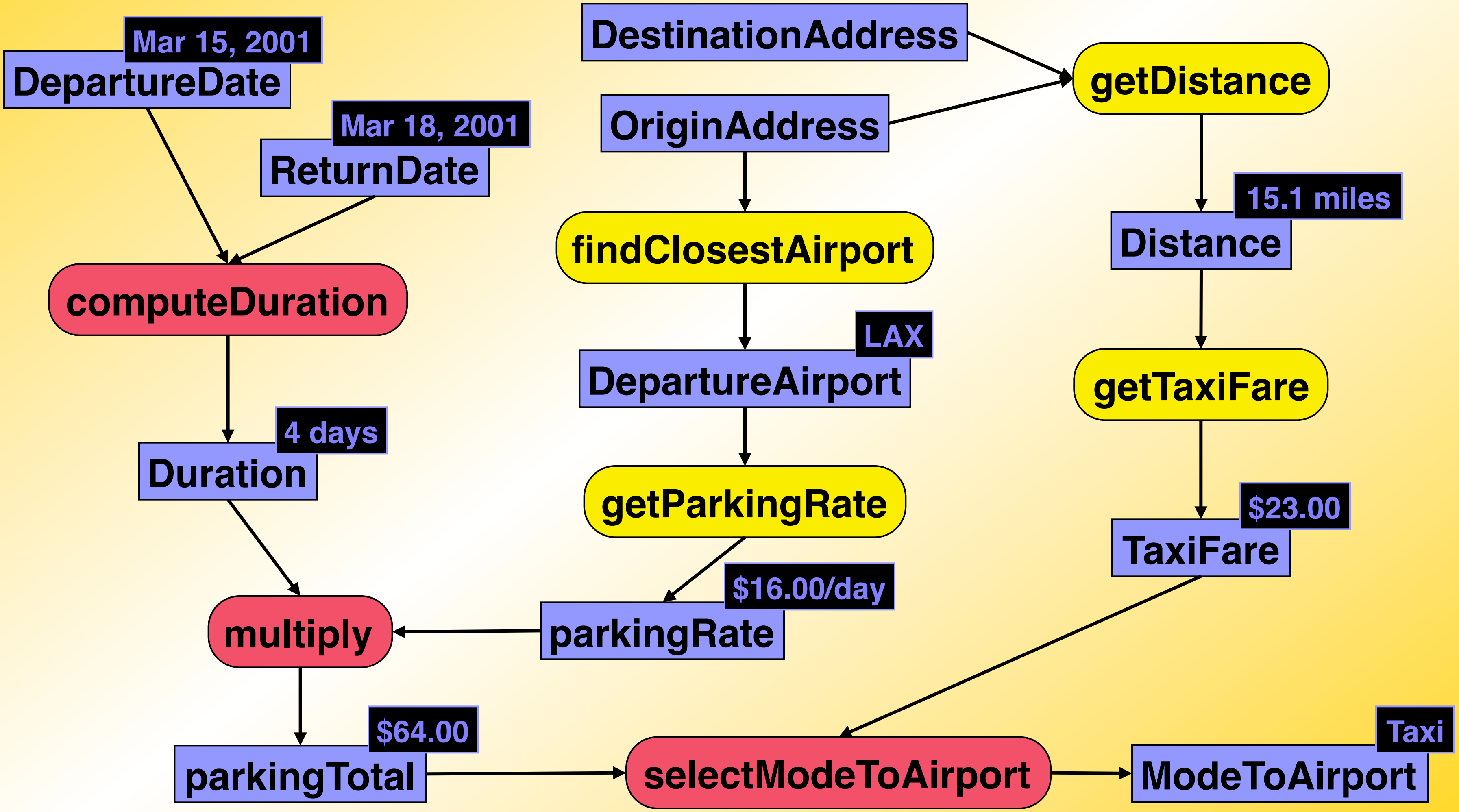
Predicted Arrival: Apr | 18 | 2002 | 08:50 PM
Month | Day | Year | Time

Taxi fare: 34.20

Total Drive: 23.3 | 0 | 28
DistYahoo | Hrs4 | Mins

Maps

Constraint Network: Drive or Taxi?



Summary

- Integration of wide range of data from many different sources
- Tight integration of data using constraints to capture the dependencies
- Supports better decision making
  - Easy to consider costs of specific choices
  - Easy to compare tradeoffs

Outline

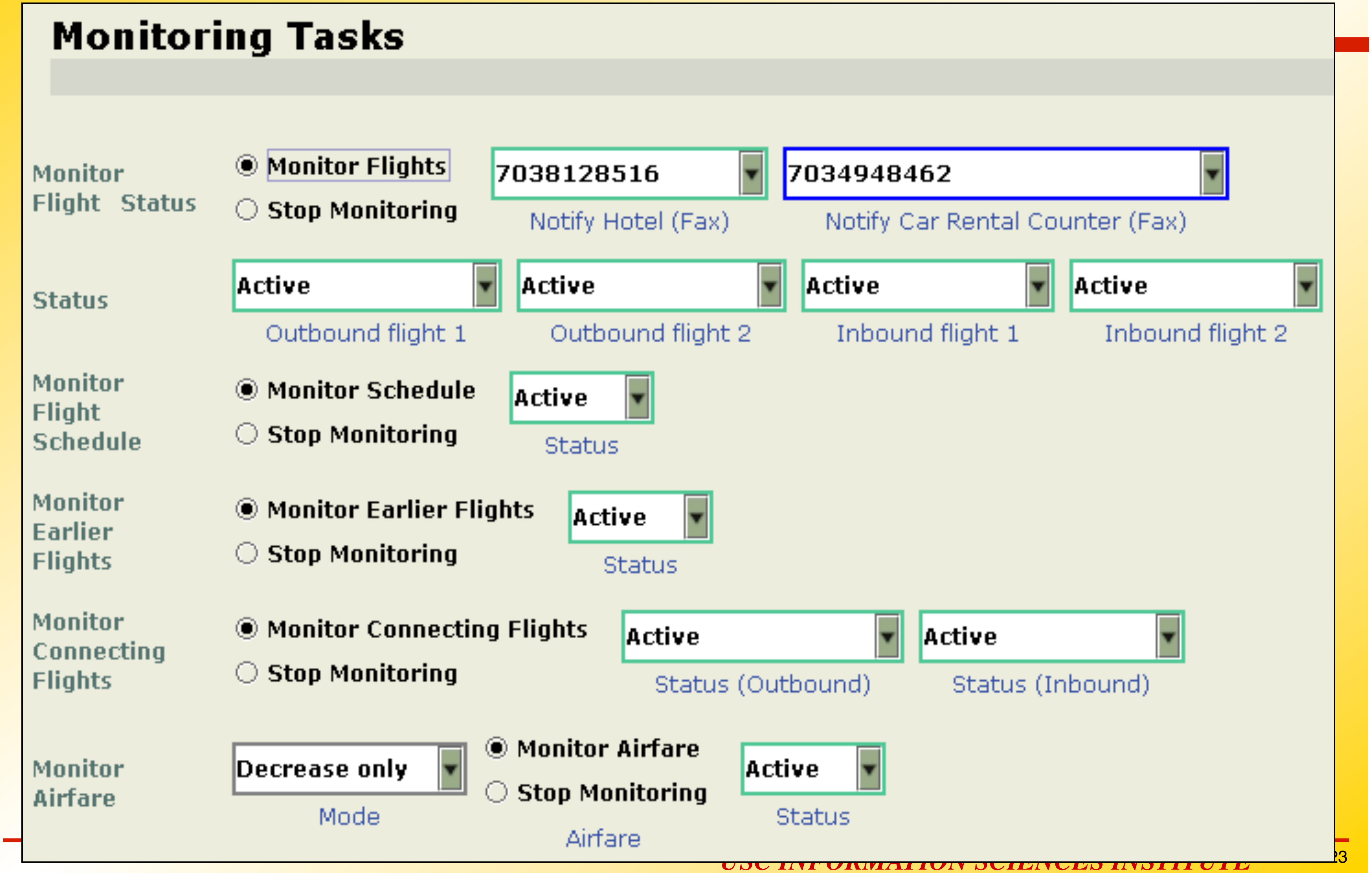
- The Electric Elves
  - Overview
- The Travel Elves
  - Constraint-based data integration
  - Efficient plan execution
- Beyond the Elves
  - Deployment to DARPA
  - Some lessons
    - Learning expectations
    - Finding alternative sources
    - Predicting prices
- Discussion

Agents for Monitoring Travel

- Many opportunities and possible problems can arise during travel
- Current environment:
  - Wide access to data
  - Abundance of computer resources
  - Availability of cell phones and portable computers
- Makes it possible to monitor all aspects of a trip
- Create personal assistants that monitor your travel plan to
  - exploit opportunities
  - avoid problems

Monitoring Travel Plans

e-Elves



Monitoring Tasks

Monitor Flight Status
- ● Monitor Flights
- ○ Stop Monitoring
- Notify Hotel (Fax): 7038128516
- Notify Car Rental Counter (Fax): 7034948462

Status
- Outbound flight 1: Active
- Outbound flight 2: Active
- Inbound flight 1: Active
- Inbound flight 2: Active

Monitor Flight Schedule
- ● Monitor Schedule
- ○ Stop Monitoring
- Status: Active

Monitor Earlier Flights
- ● Monitor Earlier Flights
- ○ Stop Monitoring
- Status: Active

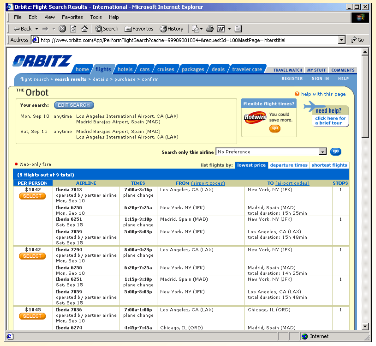
Monitor Connecting Flights
- ● Monitor Connecting Flights
- ○ Stop Monitoring
- Status (Outbound): Active
- Status (Inbound): Active

Monitor Airfare
- Mode: Decrease only
- Airfare: ● Monitor Airfare ○ Stop Monitoring
- Status: Active

Agents Deployed to Monitor Travel Itinerary

Travel Itinerary

GRID



Orbitz: Flight Search Results - International - Microsoft Internet Explorer

Address: http://www.orbitz.com/App/PerformFlightSearch?cache=99989081084&requestId=100&lastPage=interstitial

THE Orbot

help with this page

Your search: EDIT SEARCH

Mon, Sep 10 anytime Los Angeles International Airport, CA (LAX)
Madrid Barajas Airport, Spain (MAD)

Sat, Sep 15 anytime Madrid Barajas Airport, Spain (MAD)
Los Angeles International Airport, CA (LAX)

Search only this airline: No Preference

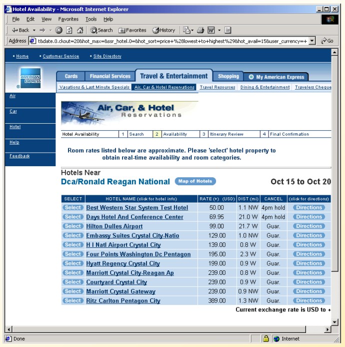
● Web-only fare

list flights by: lowest price | departure times | shortest flights

(9 flights out of 9 total)

PER PERSON	AIRLINE	TIMES	FROM (airport codes)	TO (airport codes)	STOPS
$1842 SELECT	Iberia 7013 operated by partner airline Mon, Sep 10	7:00a-3:16p plane change	Los Angeles, CA (LAX)	New York, NY (JFK)	1
	Iberia 6250 Mon, Sep 10	6:20p-7:25a	New York, NY (JFK)	Madrid, Spain (MAD) total duration: 15h 25min	
	Iberia 6251 Sat, Sep 15	1:15p-3:10p plane change	Madrid, Spain (MAD)	New York, NY (JFK)	1
	Iberia 7059 operated by partner airline Sat, Sep 15	5:00p-8:03p	New York, NY (JFK)	Los Angeles, CA (LAX) total duration: 15h 48min	
$1842 SELECT	Iberia 7294 operated by partner airline Mon, Sep 10	8:00a-4:23p plane change	Los Angeles, CA (LAX)	New York, NY (JFK)	1
	Iberia 6250 Mon, Sep 10	6:20p-7:25a	New York, NY (JFK)	Madrid, Spain (MAD) total duration: 14h 25min	
	Iberia 6251 Sat, Sep 15	1:15p-3:10p plane change	Madrid, Spain (MAD)	New York, NY (JFK)	1
	Iberia 7059 operated by partner airline Sat, Sep 15	5:00p-8:03p	New York, NY (JFK)	Los Angeles, CA (LAX) total duration: 15h 48min	
$1845 SELECT	Iberia 7036 operated by partner airline Mon, Sep 10	7:00a-1:00p plane change	Los Angeles, CA (LAX)	Chicago, IL (ORD)	1
	Iberia 6274	4:45p-7:45a	Chicago, IL (ORD)	Madrid, Spain (MAD)	

Flight Prices & Schedules



Hotel Availability - Microsoft Internet Explorer

Address: t&date.0.ckout=20&hot_max=&ssr_hotel.0=&hot_sort=price+%28lowest+to+highest%29&hot_avail=15&user_currency=+

Hotel Availability | 1 Search | 2 Availability | 3 Itinerary Review | 4 Final Confirmation

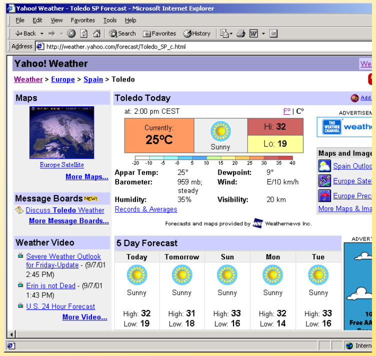
Room rates listed below are approximate. Please 'select' hotel property to obtain real-time availability and room categories.

Hotels Near
Dca/Ronald Reagan National — Map of Hotels — Oct 15 to Oct 20

SELECT	HOTEL NAME (click for hotel info)	RATE (+) (USD)	DIST (mi)	CANCEL	(click for directions)
Select	Best Western Star System Test Hotel	50.00	1.1 NW	4pm hold	Directions
Select	Days Hotel And Conference Center	69.95	21.0 W	4pm hold	Directions
Select	Hilton Dulles Airport	99.00	21.7 W	Guar.	Directions
Select	Embassy Suites Crystal City-Natio	129.00	1.0 NW	Guar.	Directions
Select	H I Natl Airport Crystal City	139.00	0.8 W	Guar.	Directions
Select	Four Points Washington Dc Pentagon	195.00	2.3 W	Guar.	Directions
Select	Hyatt Regency Crystal City	199.00	0.9 W	Guar.	Directions
Select	Marriott Crystal City-Reagan Ap	239.00	0.8 W	Guar.	Directions
Select	Courtyard Crystal City	239.00	0.9 W	Guar.	Directions
Select	Marriott Crystal Gateway	239.00	0.9 NW	Guar.	Directions
Select	Ritz Carlton Pentagon City	389.00	1.3 NW	Guar.	Directions

Current exchange rate is USD to +

Flight Status



Yahoo! Weather - Toledo SP Forecast - Microsoft Internet Explorer

Address: http://weather.yahoo.com/forecast/Toledo_SP_c.html

Yahoo! Weather

Weather > Europe > Spain > Toledo

Toledo Today

at: 2:00 pm CEST — F° | C°

Currently: 25°C — Sunny — Hi: 32 — Lo: 19

Appar Temp:	25°	Dewpoint:	9°
Barometer:	959 mb; steady	Wind:	E/10 km/h
Humidity:	35%	Visibility:	20 km

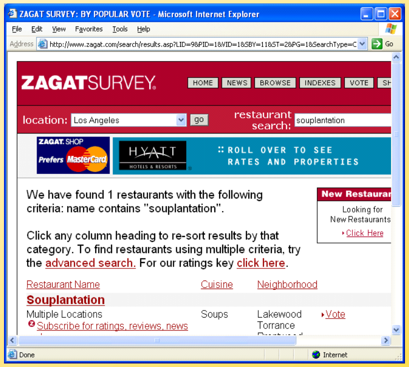
Records & Averages

Forecasts and maps provided by Weathernews Inc.

5 Day Forecast

Today	Tomorrow	Sun	Mon	Tue
Sunny	Sunny	Sunny	Sunny	Sunny
High: 32 Low: 19	High: 31 Low: 18	High: 33 Low: 16	High: 32 Low: 14	High: 33 Low: 16

Weather



ZAGAT SURVEY: BY POPULAR VOTE - Microsoft Internet Explorer

Address: http://www.zagat.com/search/results.asp?LID=9&PID=1&VID=1&SBY=11&ST=2&PG=1&SearchType=C

ZAGATSURVEY.
HOME
NEWS
BROWSE
INDEXES
VOTE
ZAGAT SHOP Prefers MasterCard
HYATT HOTELS & RESORTS
ROLL OVER TO SEE RATES AND PROPERTIES
New Restaurar
Looking for New Restaurants
Click Here

location: Los Angeles — go — restaurant search: souplantation

We have found 1 restaurants with the following criteria: name contains "souplantation".

Click any column heading to re-sort results by that category. To find restaurants using multiple criteria, try the advanced search. For our ratings key click here.

Restaurant Name	Cuisine	Neighborhood	
Souplantation			
Multiple Locations Subscribe for ratings, reviews, news	Soups	Lakewood Torrance	Vote

Restaurants

Example Agents

- Flight-Status Agent:
  - Flight delayed message:

    Your United Airlines flight 190 has been delayed.

    It was originally scheduled to depart at 11:45 AM and is now scheduled to depart at 12:30 PM.

    The new arrival time is 7:59 PM.
  - Flight cancelled message:

    Your Delta Air Lines flight 200 has been cancelled.
  - Fax to hotel message:

    Attention: Registration Desk

    I am sending this message on behalf of David Pynadath, who has a reservation at your hotel. David Pynadath is on United Airlines 190, which is now scheduled to arrive at IAD at 7:59 PM. Since the flight will be arriving late, I would like to request that you indicate this in the reservation so that the room is not given away.

Monitoring Agents

- Airfare Agent: Airfare dropped message

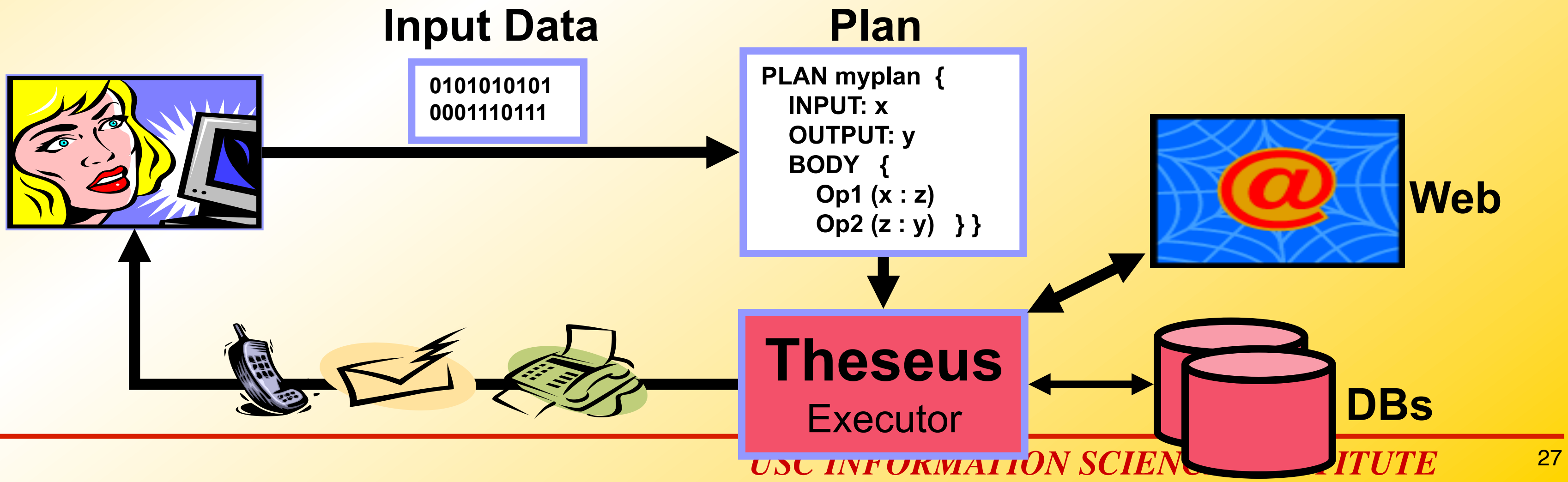
```
The airfare for your American Airlines itinerary
(IAD - LAX) dropped to $281.
```

- Earlier-Flight Agent: Earlier flights message

```
The status of your currently scheduled flight is:
# 190 LAX (11:45 AM) - IAD (7:29 PM) 45 minutes Late
If you would like to return earlier, the following
United Airlines flights will arrive earlier than your
scheduled flights:
# 946 LAX (8:31 AM) - IAD (3:35 PM) 11 minutes Late
--------
# 388 LAX (9:25 AM) - DEN (12:25 PM) 10 minutes Late
# 1534 DEN (1:20 PM) - IAD (6:06 PM) On Time
```

Theseus

- A plan language and execution system for building Web-based information-gathering and monitoring agents [Barish & Knobock, JAIR'05]
  - *Efficient* enough for near-real-time monitoring: streaming dataflow
  - *Expressive* enough for integrating a variety of sources (web sites, XML, databases, …)



```
PLAN myplan  {
   INPUT: x
   OUTPUT: y
   BODY   {
      Op1 (x : z)
      Op2 (z : y)   } }
```

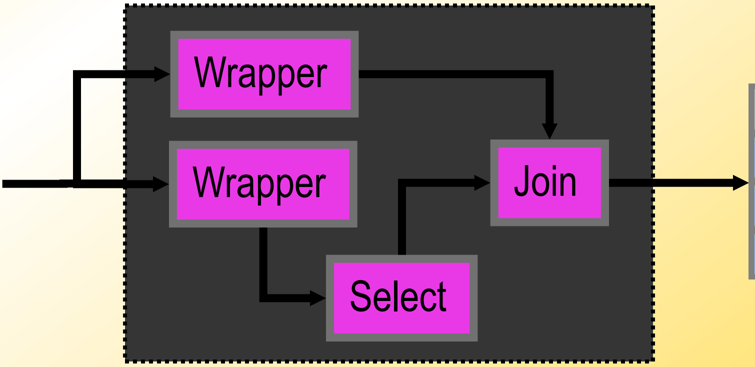
Streaming Dataflow

- Plans consist of a network of operators
  - Examples: **Wrapper**, **Select**, etc.
  - Operators produce and consume data
  - Operators "fire" upon any input data

Input relation

City	State	Max Price
Santa Monica	CA	200000

Plan



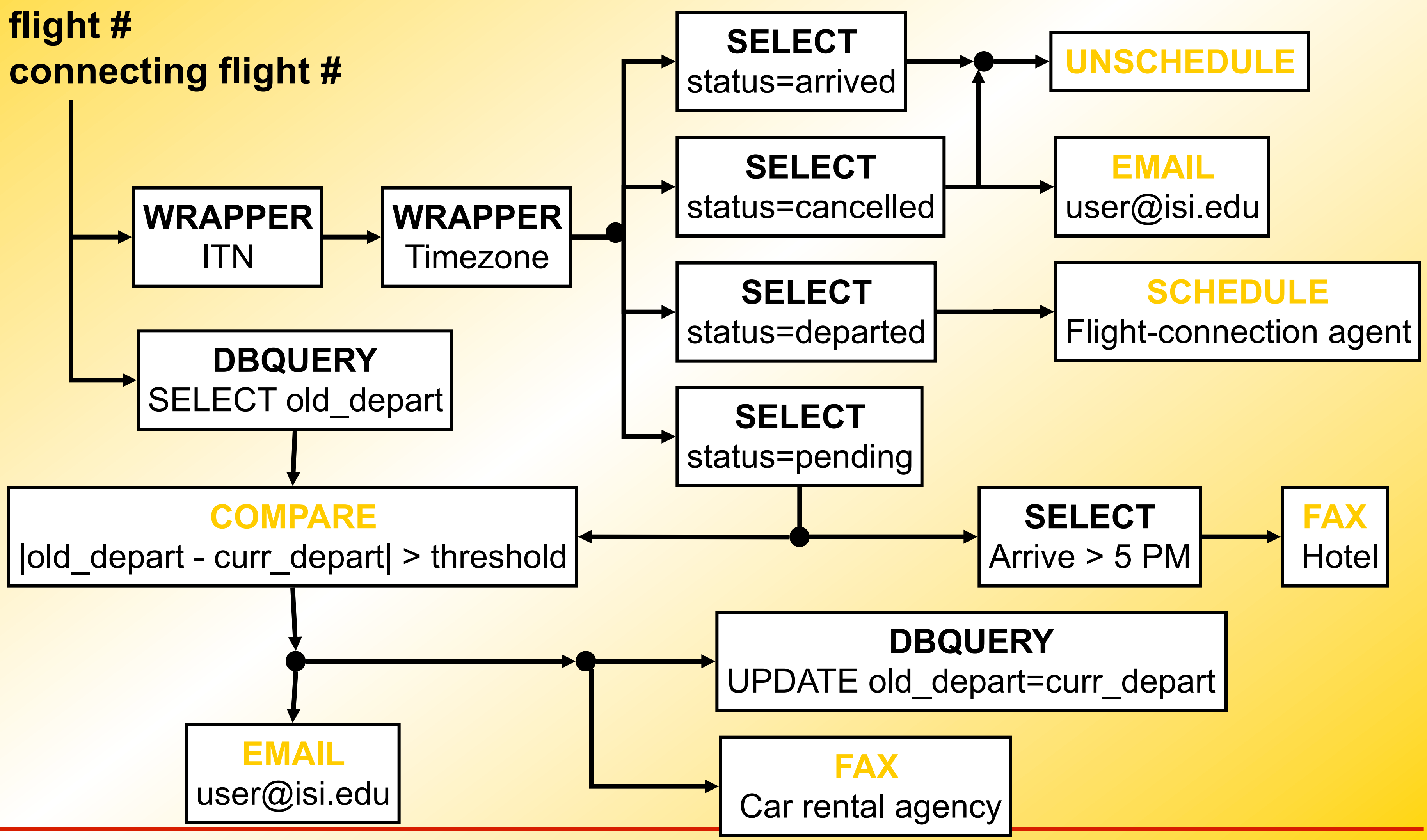
Output relation

Address
100 Main St., Santa Monica, 90292
520 4th St. Santa Monica, 90292
2 Ocean Blvd, Venice, 90292

Theseus: Efficiency Streaming Dataflow

- Dataflow-style execution
  - Operators execute when inputs become available
  - Optimizes *horizontal parallelism*
    - Plan is as parallel as its data dependencies allow
- Data Streaming
  - Data in the system represented as *relations*
    - Producer operators pipeline *tuples* to consumers
  - Optimizes *vertical parallelism*
    - Multiple operators can work on same relation concurrently

Theseus Monitoring Agent: Flight Status Plan



Outline

- The Electric Elves
  - Overview
- The Travel Elves
  - Constraint-based data integration
  - Efficient plan execution
- Beyond the Elves
  - Deployment to DARPA
  - Some lessons
    - Learning expectations
    - Finding alternative sources
    - Predicting prices
- Discussion

Deploying the Elves at DARPA

- The office elves were successful and received a lot of press
- DARPA decided they wanted a version running in their environment
- But there were issues
  - Integration into the DARPA environment
  - Privacy concerns
  - Software integration
  - …

The Travel Elves

- So we created the Travel Elves
  - No integration into the DARPA environment
  - Could run on ISI servers
  - Communicate via PDAs
- Trained the DARPA travel person to enter travel data using Heracles
- System deployed…
  - Used by program managers and office directors
  - …for a while

What Went Wrong and Why?

- Agent failures due to problems with the data
  - “Why didn’t I receive any updates from the Elves?”
  - Source changed, unexpected data, source unavailable
- Agents failures due to problems in the logic
  - “Why did the message arrive after my flight left?”
  - Time zone exception
- Unexpected behaviors
  - “Why are the Elves pestering me with price changes?”
  - Frequent price changes
- Getting the definition “right”
  - “Those !*?&%* Elves sent a message to my cell phone at 3am to tell me my 8am flight was delayed!”
  - Agent worked as designed…
- Personalized agents
  - “Could you build an agent to do <X>?”
  - Every user wants something different

Outline

- The Electric Elves
  - Overview
- The Travel Elves
  - Constraint-based data integration
  - Efficient plan execution
- Beyond the Elves
  - Deployment to DARPA
  - Some lessons
    - Learning expectations
    - Finding alternative sources
    - Predicting prices
- Discussion

Learning Data Prototypes [Lerman et al., JAIR 2003]

- Approach to learning the structure of data
- Token-level syntactic description
  - descriptive but compact
  - computationally efficient
- Structure is described by a sequence (pattern) of general and specific tokens.
- Data prototype = starting & ending patterns

STREET_ADDRESS

220 Lincoln Blvd

420 S Fairview Ave

2040 Sawtelle Blvd

start with:

_NUM _CAPS

end with:

_CAPS Blvd

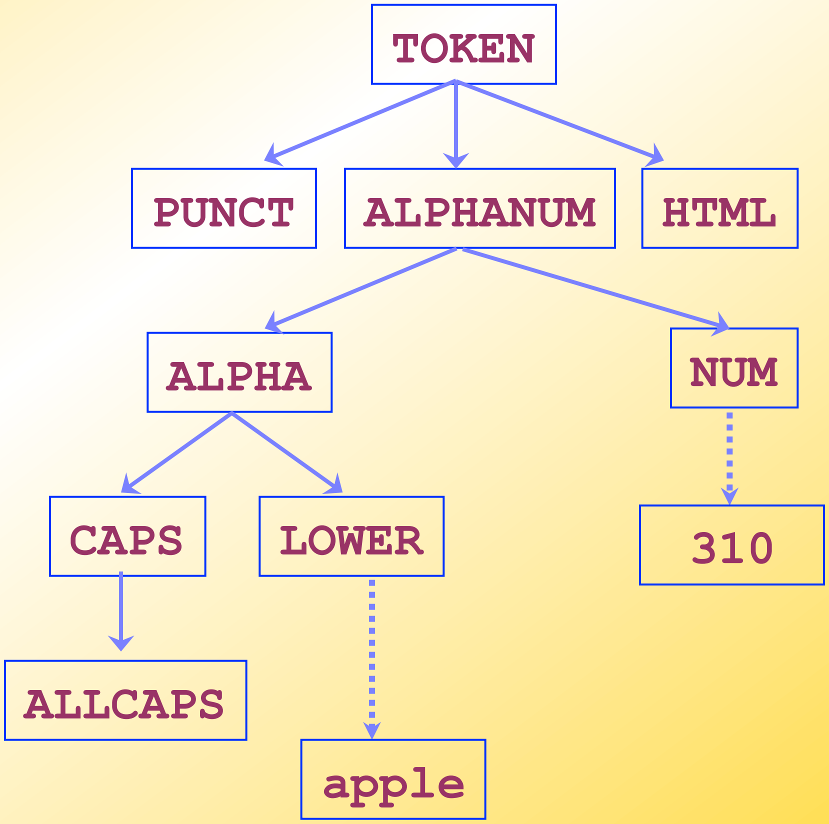
_CAPS _CAPS

Token Syntactic Hierarchy

- Tokens = words
- Syntactic types

  e.g., NUMBER, ALPHA
- Hierarchy of types

  allows generalization
- Extensible
  - new types
  - domain-specific information



Prototype Learning Algorithm

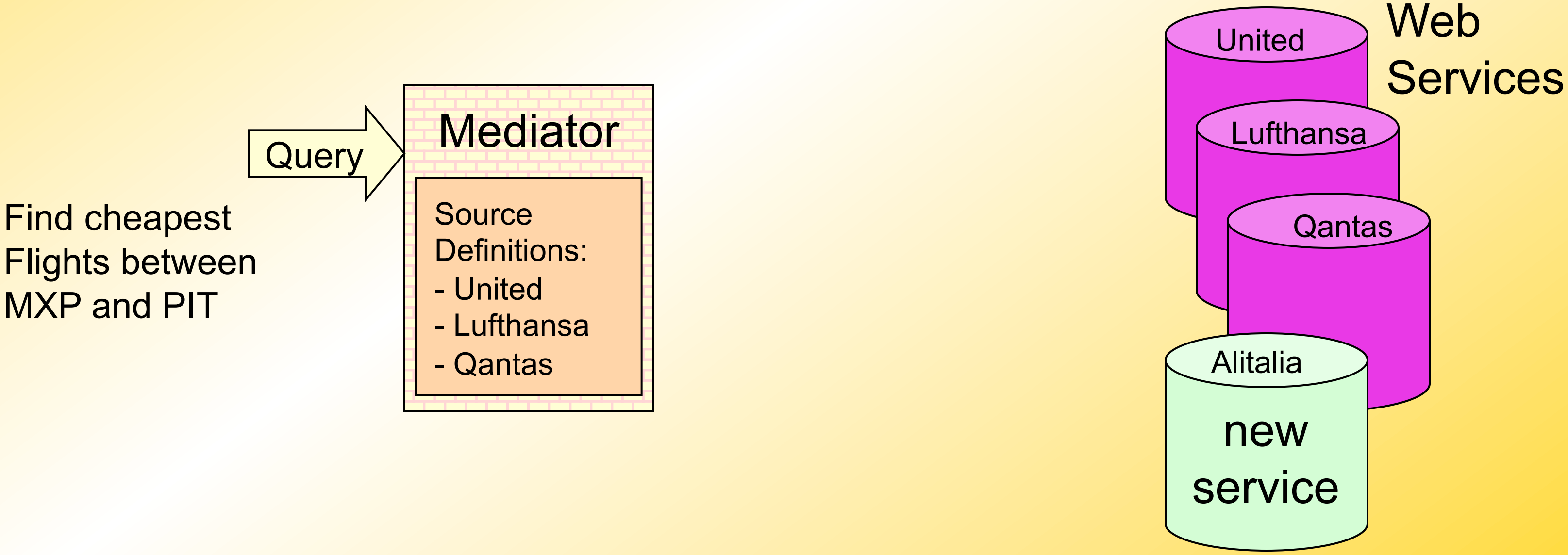
- No explicit negative examples
- Learn from positive examples of data
- Find patterns that
  - describe many of the positive examples of data
  - highly unlikely to describe a random token sequence (implicit negative examples)
- are statistically significant patterns

  at α=0.05 significance level
- **DataPro** – efficient (greedy) algorithm

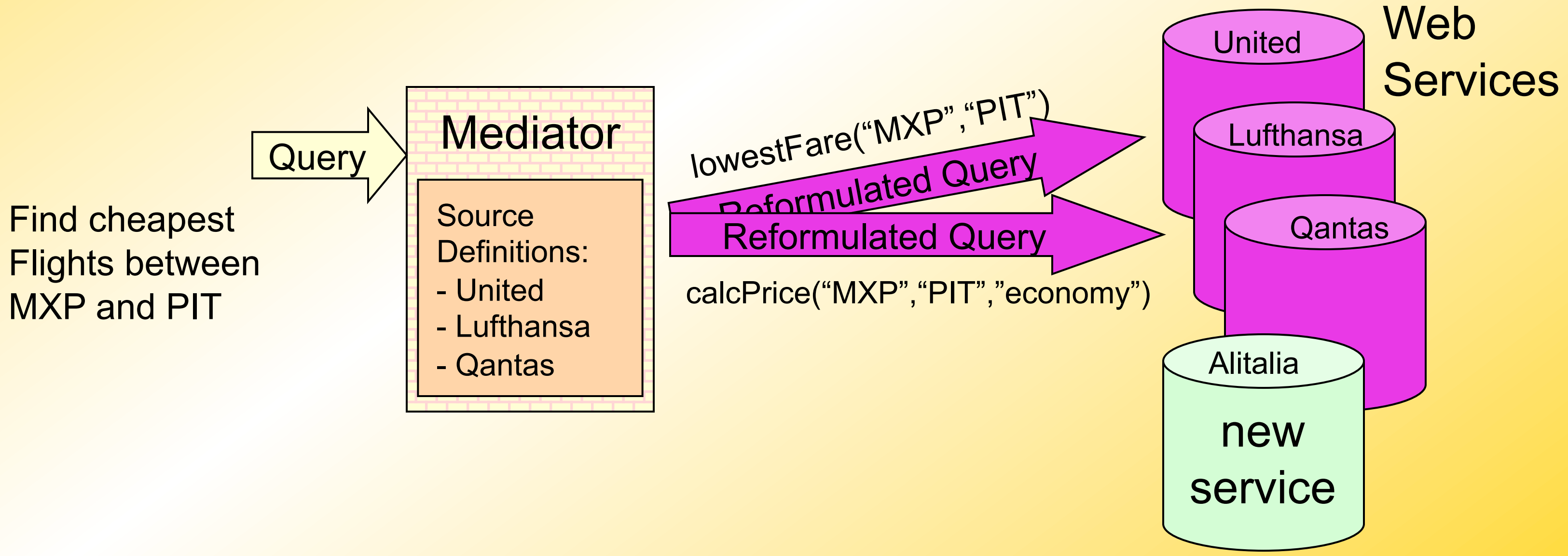
Outline

- The Electric Elves
  - ➢ Overview
- The Travel Elves
  - ➢ Constraint-based data integration
  - ➢ Efficient plan execution
- Beyond the Elves
  - ➢ Deployment to DARPA
  - ➢ Some lessons
    - – Learning expectations
    - – Finding alternative sources
    - – Predicting prices
- Discussion

Learning New Sources of Data [Carman & Knoblock, 2005]

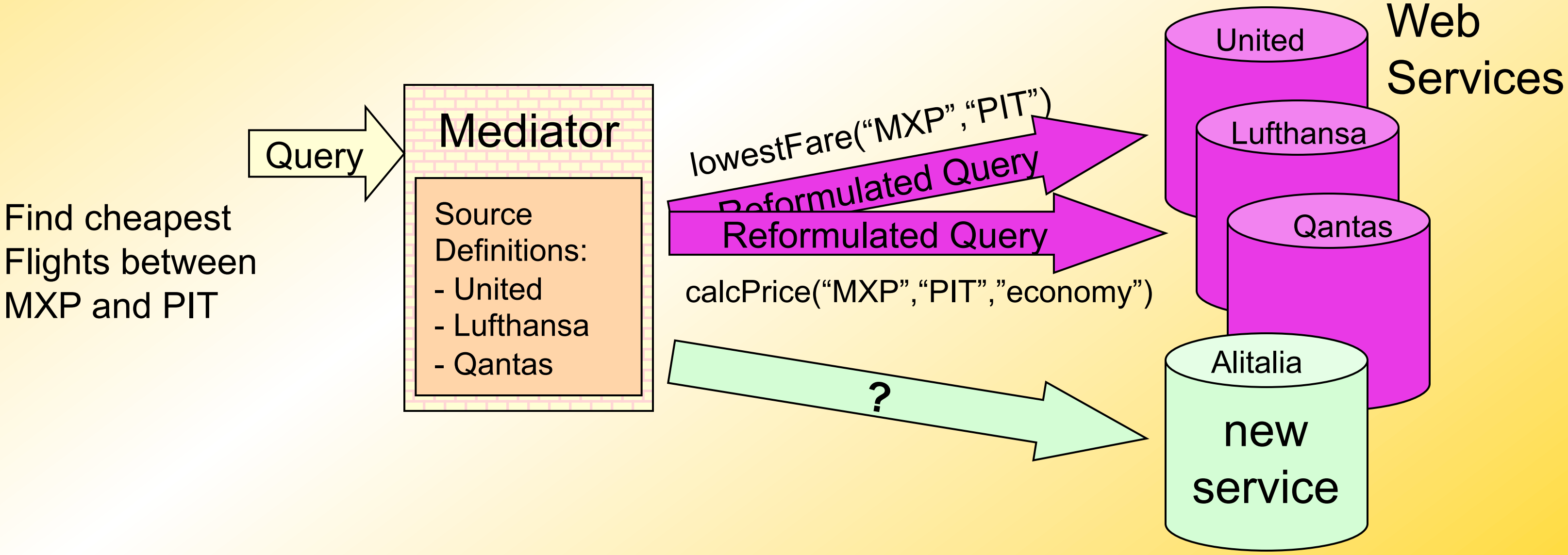


Learning New Sources of Data [Carman & Knoblock, 2005]



Learning New Sources of Data [Carman & Knoblock, 2005]

- Need source definitions to incorporate new data
- Time consuming and difficult to write these descriptions
- Can we discover definitions automatically?



The Framework

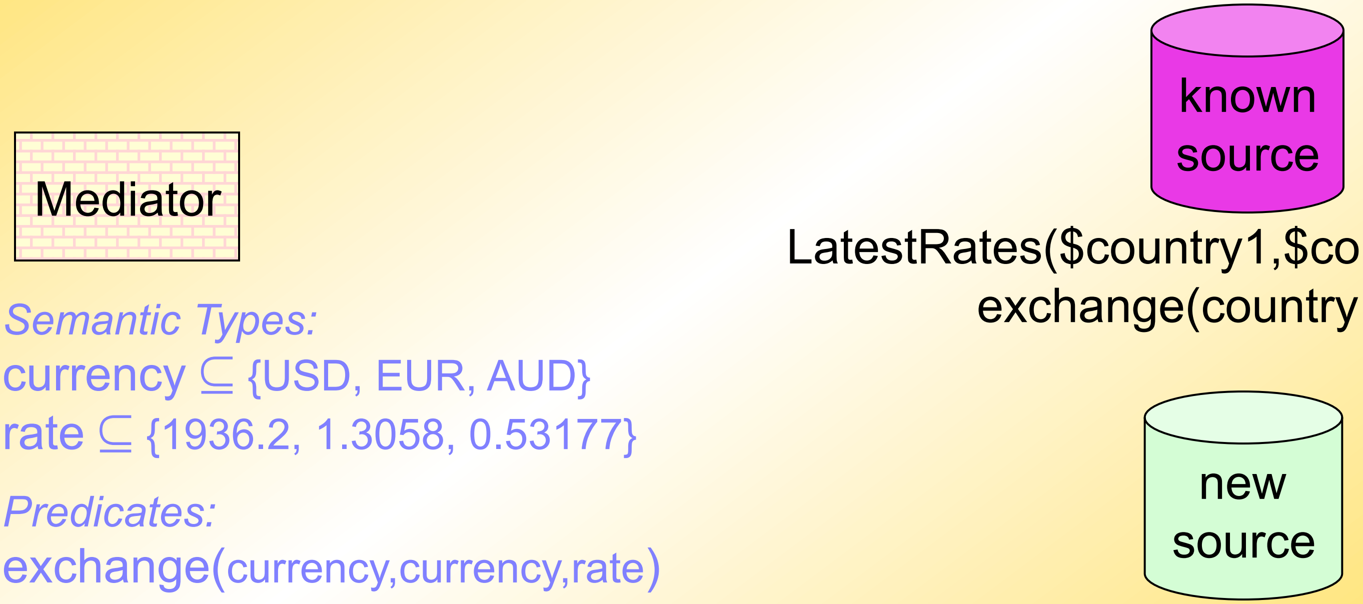
Intuition: *Services often have similar semantics, so we should be able to use what we know to induce that which we don't*

Two phase algorithm

For each operation provided by the new service:

1. Classify its input/output data types
   - *Classify* inputs based on metadata similarity
   - *Invoke* operation & classify outputs based on data
2. Induce a source definition
   - *Generate* candidates via Inductive Logic Programming
   - *Test* individual candidates by reformulating them

Inducing Source Definitions: A Simple Example



Mediator

Semantic Types:

currency ⊆ {USD, EUR, AUD}

rate ⊆ {1936.2, 1.3058, 0.53177}

Predicates:

exchange(currency,currency,rate)

known source

LatestRates($country1,$country2,rate):-
exchange(country1,country2,rate)

new source

RateFinder($fromCountry,$toCountry,val):- ?

Inducing Source Definitions: A Simple Example

- Step 1: use metadata to classify input types

Mediator

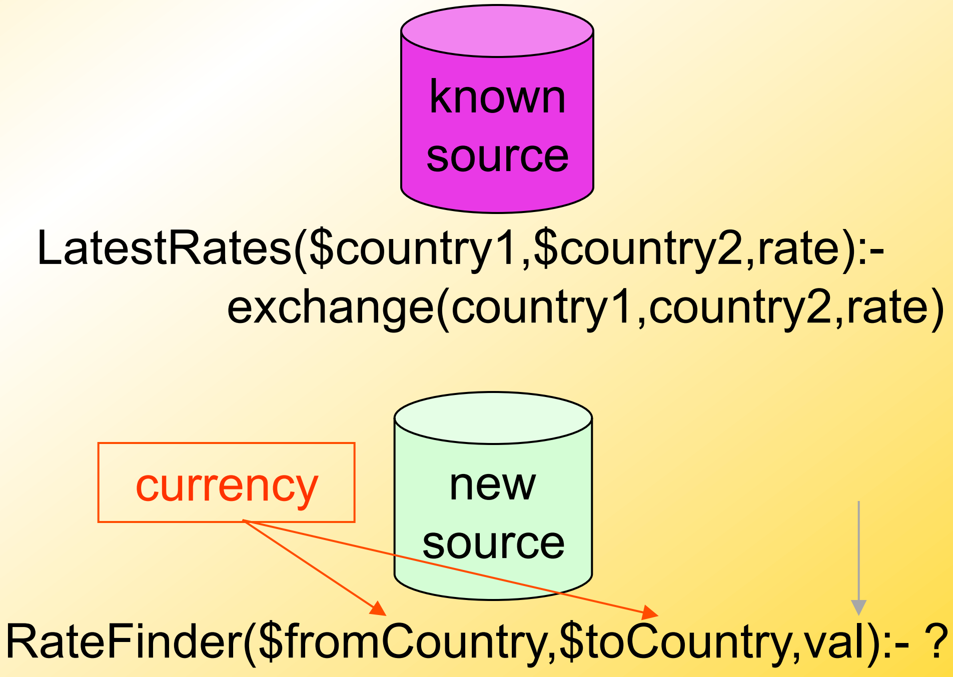
Semantic Types:

currency ⊆ {USD, EUR, AUD}

rate ⊆ {1936.2, 1.3058, 0.53177}

Predicates:

exchange(currency,currency,rate)



LatestRates($country1,$country2,rate):-
exchange(country1,country2,rate)

RateFinder($fromCountry,$toCountry,val):- ?

Inducing Source Definitions: A Simple Example

- Step 1: use metadata to classify input types
- Step 2: invoke service and classify output types

Mediator

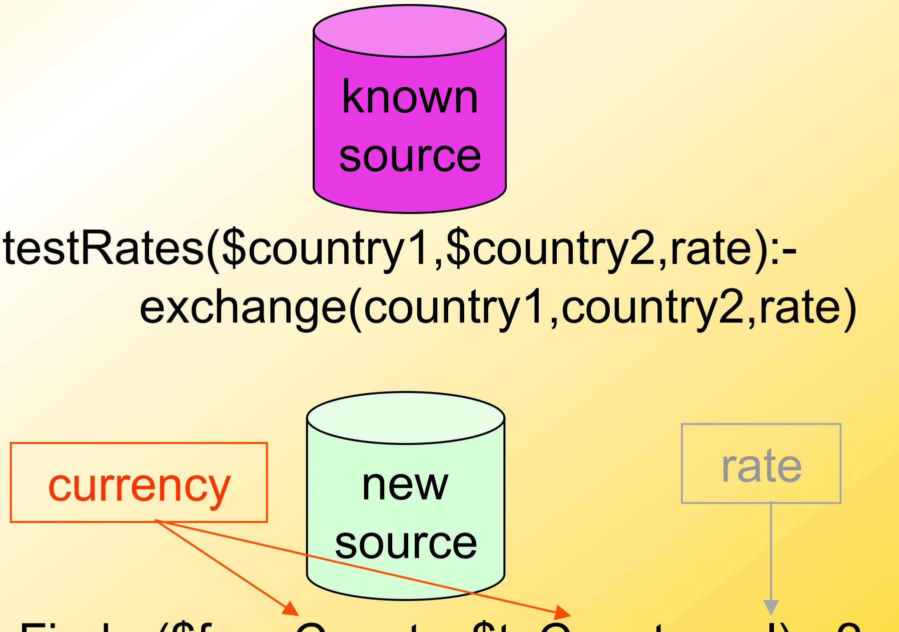
Semantic Types:

currency ⊆ {USD, EUR, AUD}

rate ⊆ {1936.2, 1.3058, 0.53177}

Predicates:

exchange(currency,currency,rate)



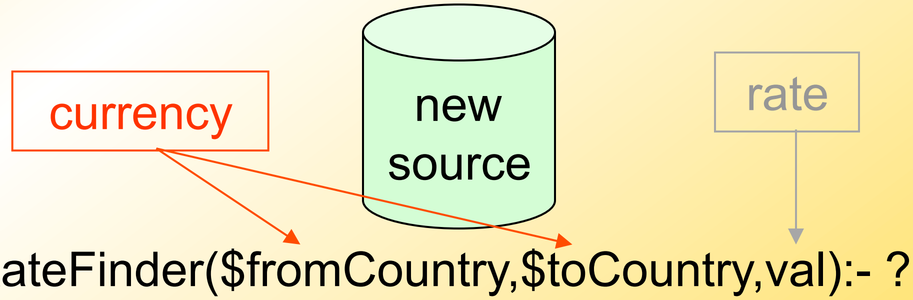
LatestRates($country1,$country2,rate):-
exchange(country1,country2,rate)

RateFinder($fromCountry,$toCountry,val):- ?

{<EUR,USD,1.30799>,<USD,EUR,0.764526>,…}

Inducing Source Definitions: A Simple Example

- Step 3: generate plausible source definitions



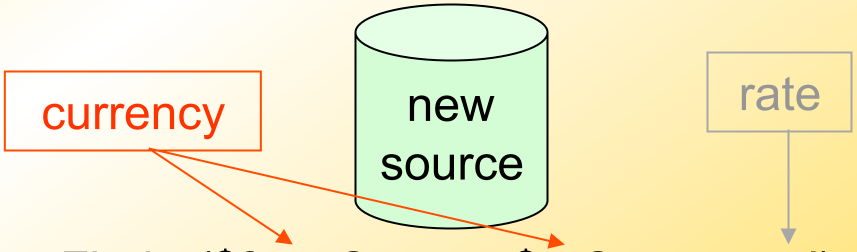
RateFinder($fromCountry,$toCountry,val):- ?

Mediator

Predicates:
exchange(currency,currency,rate)

Inducing Source Definitions: A Simple Example

- Step 3: generate plausible source definitions



RateFinder($fromCountry,$toCountry,val):- ?

def_1($from, $to, val) :- exchange(from,to,val)

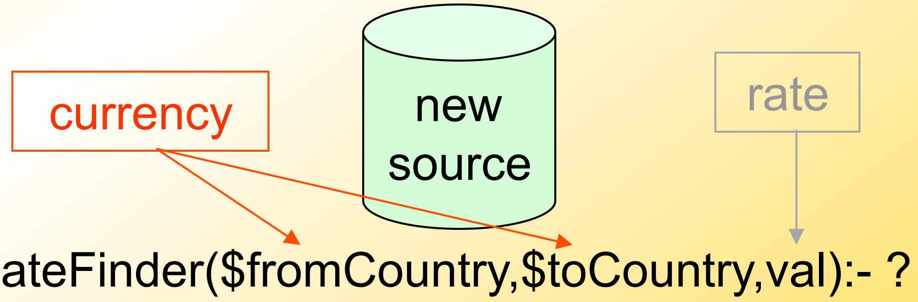
def_2($from, $to, val) :- exchange(to,from,val)

Mediator

Predicates:
exchange(currency,currency,rate)

Inducing Source Definitions: A Simple Example

- Step 3: generate plausible source definitions
- Step 4: reformulate in terms of other sources



RateFinder($fromCountry,$toCountry,val):- ?

Mediator

Predicates:
exchange(currency,currency,rate)

def_1($from, $to, val) :- exchange(from,to,val)

def_2($from, $to, val) :- exchange(to,from,val)

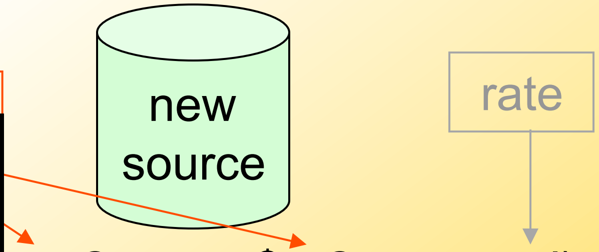
def_1($from, $to, val) :- LatestRates(from,to,val)

def_2($from, $to, val) :- LatestRates(to,from,val)

Inducing Source Definitions: A Simple Example

- Step 3: generate plausible source definitions
- Step 4: reformulate in terms of other sources
- Step 5: invoke service and compare output

Input	RateFinder	Def_1	Def_2
<EUR,USD>	1.30799	1.30772	0.764692
<USD,EUR>	0.764526	0.764692	1.30772
<EUR,AUD>	1.68665	1.68979	0.591789



romCountry,$toCountry,val):- ?

Mediator

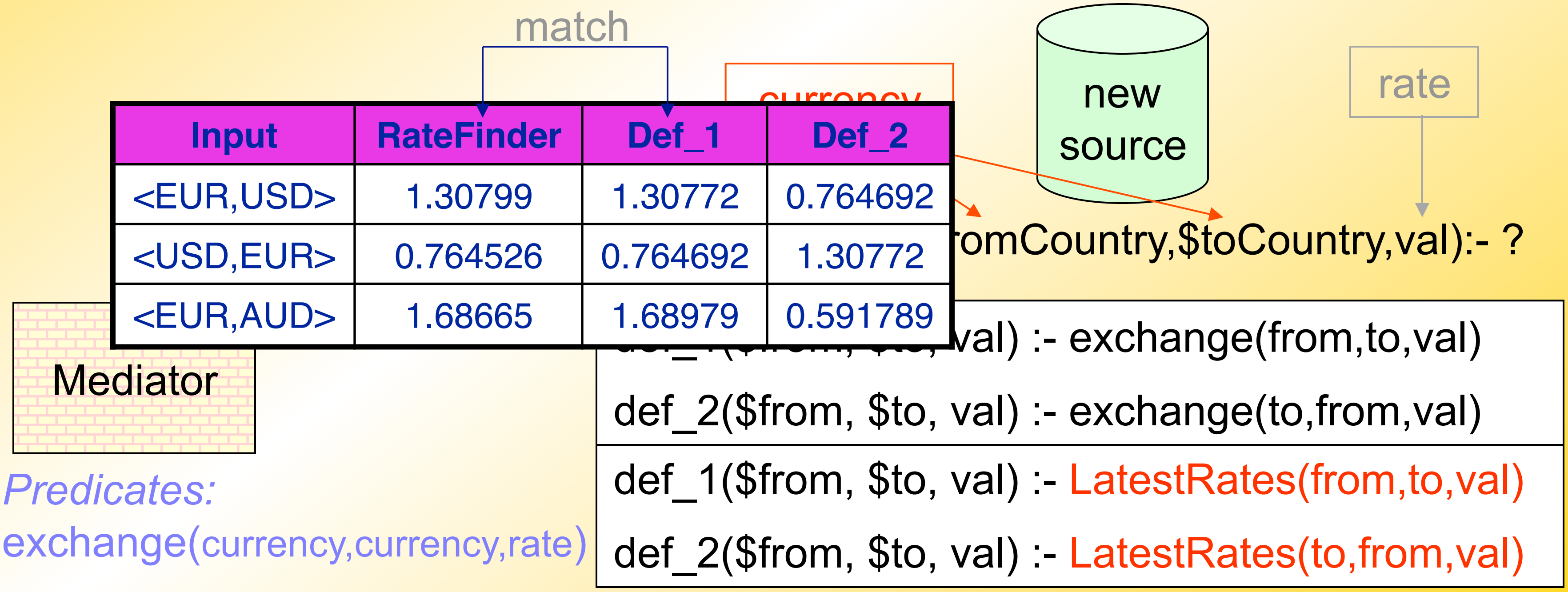
Predicates:
exchange(currency,currency,rate)

def_1($from, $to, val) :- exchange(from,to,val)
def_2($from, $to, val) :- exchange(to,from,val)

def_1($from, $to, val) :- LatestRates(from,to,val)
def_2($from, $to, val) :- LatestRates(to,from,val)

Inducing Source Definitions: A Simple Example

- Step 3: generate plausible source definitions
- Step 4: reformulate in terms of other sources
- Step 5: invoke service and compare output



Input	RateFinder	Def_1	Def_2
<EUR,USD>	1.30799	1.30772	0.764692
<USD,EUR>	0.764526	0.764692	1.30772
<EUR,AUD>	1.68665	1.68979	0.591789

Predicates:
exchange(currency,currency,rate)

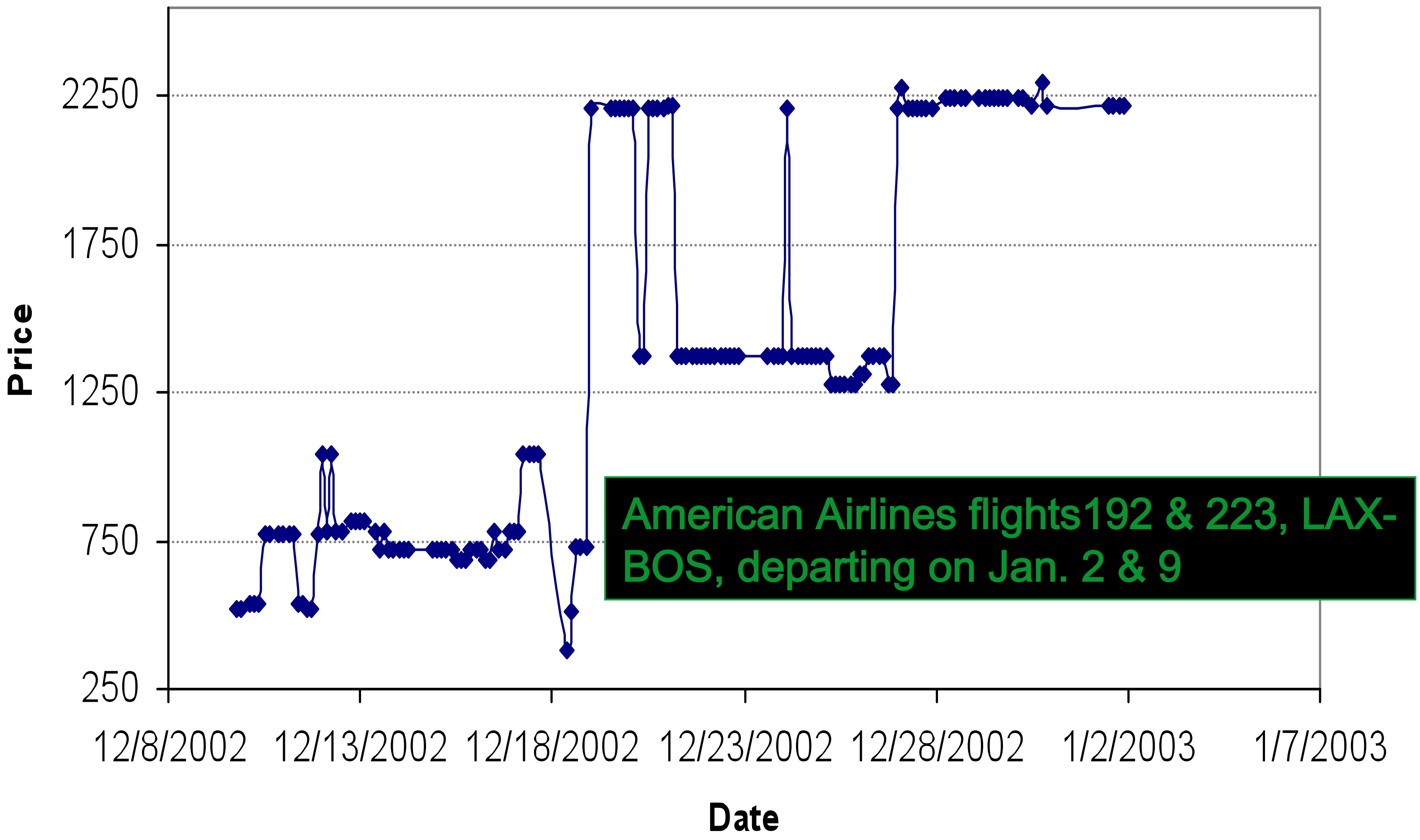
def_1($from, $to, val) :- exchange(from,to,val)
def_2($from, $to, val) :- exchange(to,from,val)

def_1($from, $to, val) :- LatestRates(from,to,val)
def_2($from, $to, val) :- LatestRates(to,from,val)

Outline

- The Electric Elves
  - Overview
- The Travel Elves
  - Constraint-based data integration
  - Efficient plan execution
- Beyond the Elves
  - Deployment to DARPA
  - Some lessons
    - Learning expectations
    - Finding alternative sources
    - Predicting prices
- Discussion

Predicting Airline Prices [Etzioni et al., KDD 2003]



American Airlines flights192 & 223, LAX-BOS, departing on Jan. 2 & 9

Hamlet: To Buy or Not to Buy

- Collected airline flight data over several months
- Developed a learning algorithm to predict whether to buy immediately or wait to buy a ticket
- Exploits the fact that airline pricing is done with a relatively static, but unknown algorithm
- Pricing can be learned by considering the pricing on the same flight on previous days

Data Set

- Extracted data from online sources using wrappers
- Collected over 12,000 price observations:
  - Lowest available fare for a one-week roundtrip
  - LAX-BOS and SEA-IAD
  - 6 airlines including American, United, etc.
  - 21 days before each flight, every 3 hours

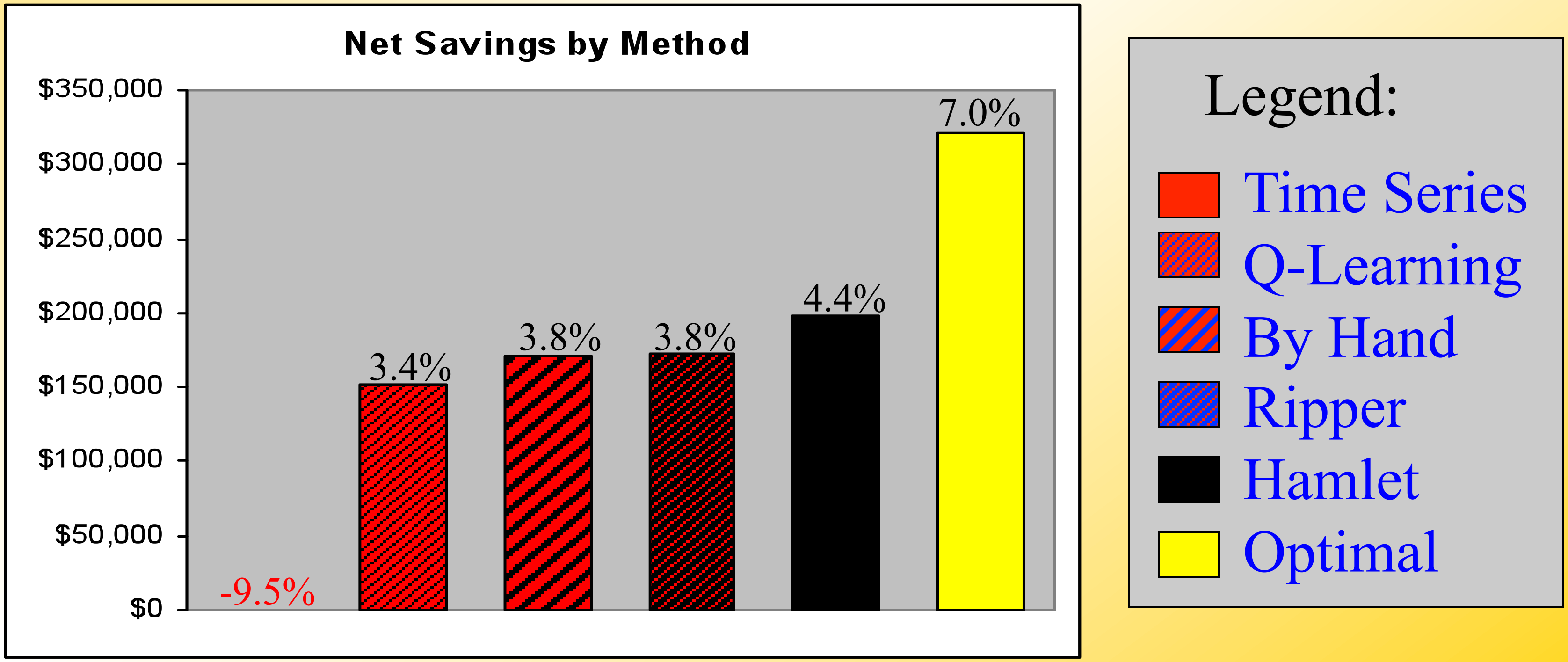
Learning Algorithm

- Stacking with three base learners:
  1. Rule learning (Ripper) (e.g., R=wait)
  2. Time series
  3. Q-learning (e.g., Q=buy)
- Ripper used as the meta-level learner.
- Output: classifies each decision point as

  ‘buy’ or ‘wait’.

Savings by Method

- **Net** savings = cost now – cost at purchase point.
- Penalty for sell out = upgrade cost. 0.42% of the time.
- Total ticket cost is $4,579,600.



Status

- Patent granted on preditive pricing of air fares based on historical data
- Technology licensed to Hamlet, Inc. started by Oren Etzioni
- Company raised $7M in VC funding and is now called Farecast.
- Stay tuned…

Discussion

- We need to build agents that:
  - Robustly accomplish their tasks, responding appropriately to failures
  - Learn from their past experience
  - Rapidly build personalized agents without manual programming
  - Communicate flexibly with humans and software agents
  - Explain their behavior both on success or failure
  - Dynamically compose new agents and behaviors from existing agents