Management Analytics

# Generalizations to Travel Time Consistency Checking

2016-01-04
IFIP TC11 - Delhi, India

Dr. Don Cohen
Dr. Fred Cohen

## What's this paper all about?

- Admissible → Reliable & Authentic
- How do we show R&A
  - to within the standard of proof?

- Like all scientific analysis:
  - Theory of the case → Testable Hypothesis
  - Test testable hypothesis by refutation

- This analysis method is about consistency checking for records involving
  - (atomic) objects
  - at (bounded) places
  - at (intervals of) times

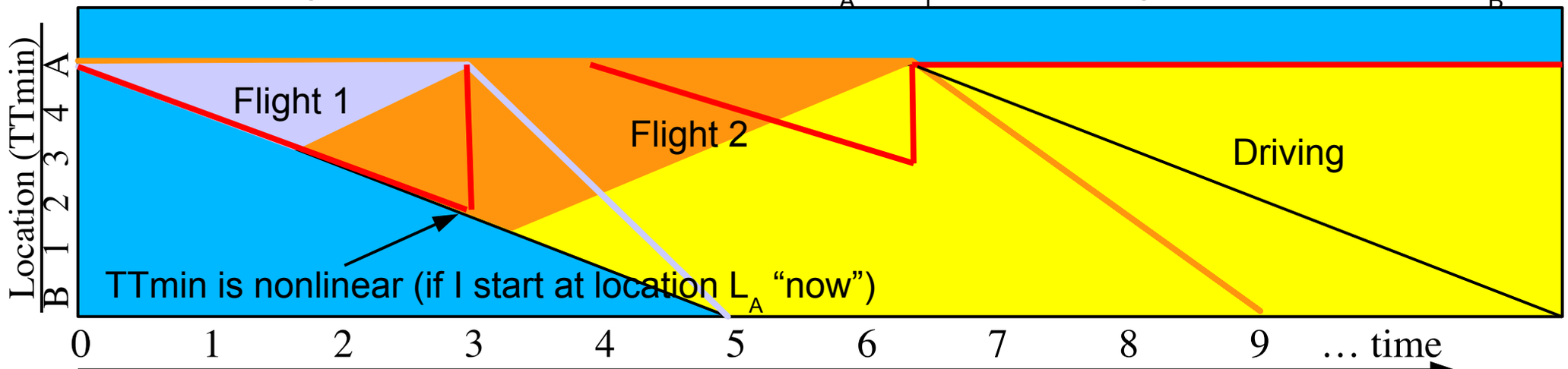
- Inconsistency of records implies
  - an atomic object was in two places at the same time

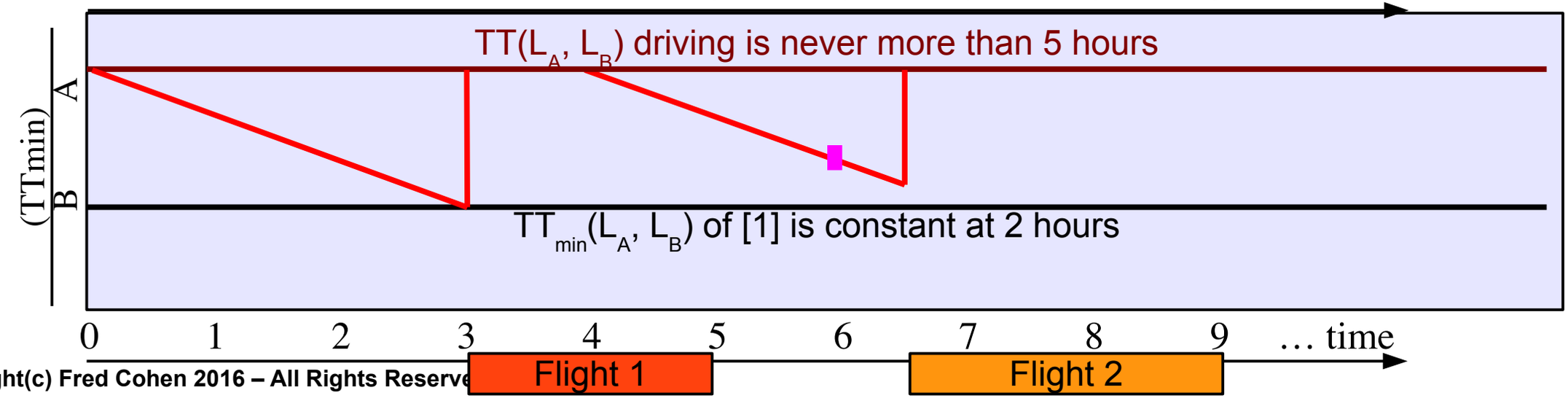
Management Analytics

# Prior Art [1]

**Time and Space Interval Record Schedule Consistency Analysis for Atomic Items without Interactions in Open Spaces with Stationary Locations**

- **Inconsistency of records implies either: (1)** An atomic object was in two places at the same time OR The records are not reliable and accurate
- **Simple example**: If you are recorded at location $L_A$ at $t_1$, when can you be at location $L_B$?

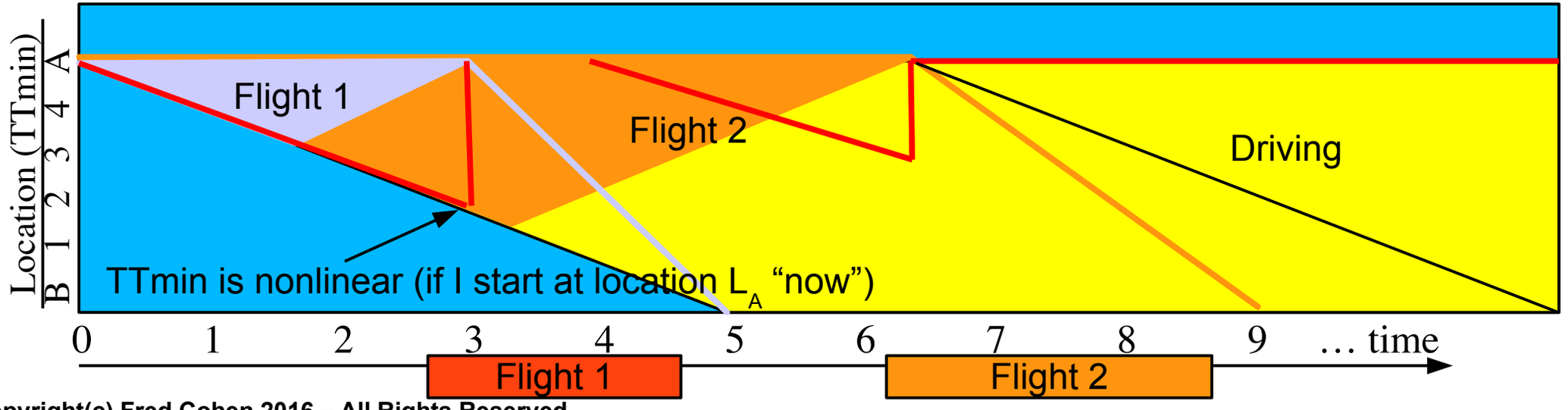


Management Analytics

# Outline

- Admissible → Reliable and Authentic
- Theory of the case → Testable Hypothesis
- Consistency checking for records involving
  - (atomic) objects
  - at (bounded) places
  - at (intervals of) times
- An atomic object in two places at once? Use travel time (TT)!



Management Analytics

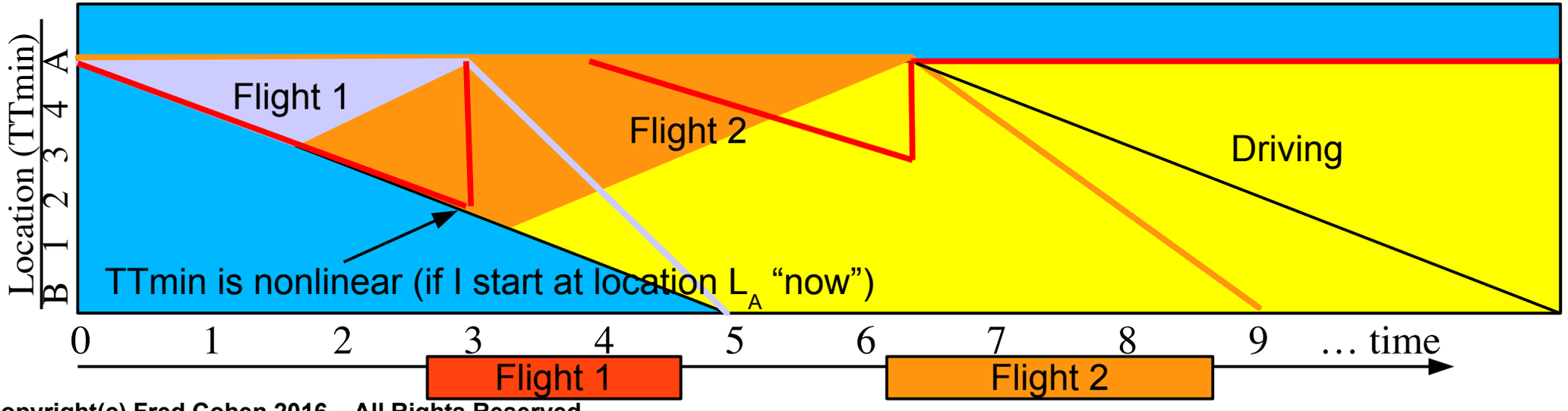
# Admissible → Reliable and Authentic

- To within the standard of proof
  - Of the nature and jurisdiction of the legal proceeding
- Admissible evidence must be shown by the party propounding it
  - To be reliable
    - The record is a true statement of fact
  - To be authentic
    - The record is what it purports to be
- Authentication (not authenticity)
  - Declaration of authenticity made by competent party, assessed based on preservation of identity and integrity over time
- Reliability
  - Declaration of reliability made by a competent party, assessed based on completeness, controls of its creation

Management Analytics

# Outline

- Admissible → Reliable and Authentic
- Theory of the case → Testable Hypothesis
- Consistency checking for records involving
  - (atomic) objects
  - at (bounded) places
  - at (intervals of) times
- An atomic object in two places at once? Use travel time (TT)!



Management Analytics

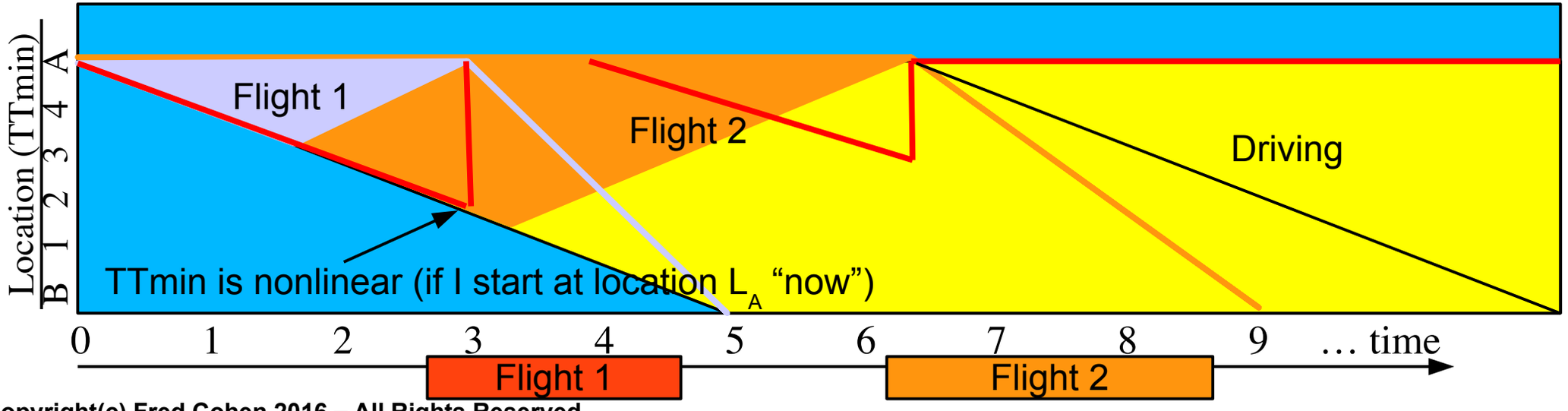
# Theory of the case

- John broke the glass at the bookstore
  - Amy says she wiped glass clean just before leaving
  - Amy says the glass was not broken when she left
  - Amy says she turned on the alarm when she left.
  - Records show the alarm was turned on at 6:59, March 12, 2015
  - Records show the alarm sounded off at 7:00 the same day
  - Police reports arrival at 7:10 the same day, found the glass broken, nobody else was present
  - Finger prints show the broken glass had John's finger prints on it
- Hypothesis to support the theory (testable hypothesis?):
  - John was at the bookstore at or about 7 PM - 7:10 that day
- How can we test this hypothesis?

Management Analytics

# Outline

- Admissible → Reliable and Authentic
- Theory of the case → Testable Hypothesis
- Consistency checking for records involving
  - (atomic) objects
  - at (bounded) places
  - at (intervals of) times
- An atomic object in two places at once? Use travel time (TT)!



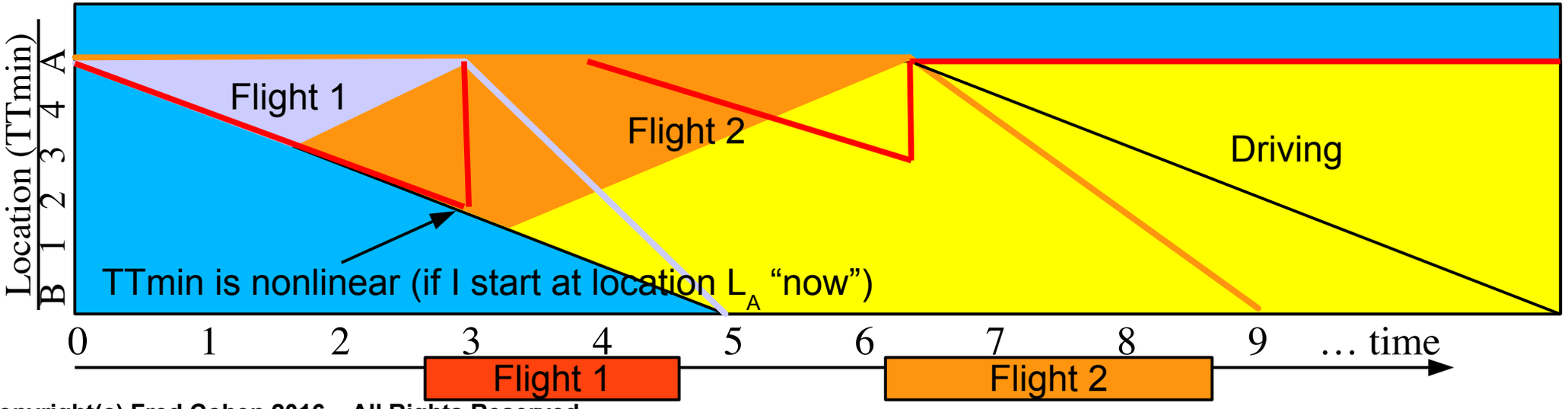
Management Analytics

# Checking the consistency of records

- If the hypothesis is true, all relevant records should be consistent with this hypothesis
- If all relevant records are not consistent with this hypothesis
  - The records are not reliable
  - The records are not authentic
  - The hypothesis is not true
- Legally (theoretically):
  - The records are not reliable → the records are not admissible
  - The records are not authentic → the records are not admissible
  - The hypothesis is not true → John wasn't there (acquittal likely)
- The remaining case may be inadequate to go to court or unconvincing to the trier of fact

# How do we check the consistency of records?

- For certain types of records, we have a way:
  - Theory: John was at the bookstore at or about 7 PM – 7:10
  - Records:
    - Those listed above for the case PLUS
    - All other records recoverable and admissible
      - Credit card usage by John
      - Records form CCTV and traffic cameras
      - GPS records form John's phone
      - Eye witness testimony (he was in my apartment at 7:30)



Management Analytics

# Methodology

- Encode all records in terms of
  - Location and location uncertainty (interval)
  - Time and time uncertainty (interval)
  - Identify that John is an “atomic object”
    - Cannot be in two distinct places at the same time
- Encode minimum travel times
  - From/to recorded locations as a function of time
  - For the relevant time frames
- Run the algorithm
  - Inconsistencies are identified
    - Sets of records that cannot all be true
    - The reasons they cannot all be true

Management Analytics

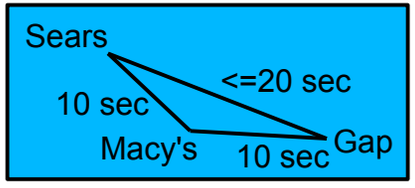
# The algorithm (note CTT)

I. group scheduling constraints by atomic object
II. for each atomic object, O, process the scheduling constraints for O:
  A. preprocess the set of scheduling constraints:
    alternate between forward and backward in time passes until
    a pass through the set makes no changes
      1. sort by start time, compare each scheduling constraint to
        those after it until time difference > max of min TT
      2. CTT:
        in cases where times require adjacent appearances, use CTT instead of TT
  B. separate the set of scheduling constraints into independent subsets
      1. sort by time, find gaps of inactivity > max of min TT
      2. CTT:
        determine whether adjacent sets must, must not or might satisfy CTT
        must not => inconsistent
        must => leave the two separated
        might => keep them joined together
  C. search for a consistent schedule using depth first search; expand-node does:
      1. find remaining scheduling constraints that could be next
        on the basis of start/end times
      2. for each candidate (until success or resource limit) create a new node
        a. CTT:
          reject the node if it violates CTT (for original algorithm CTT = TT)
        b. nonΔ≠ TT:
          if TT is nonΔ≠ then also check TT against earlier
          appearances back to max of min TT

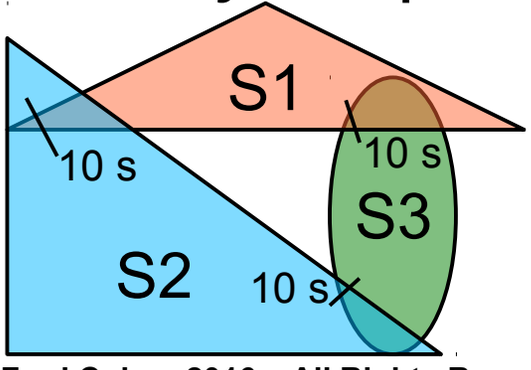
Management Analytics

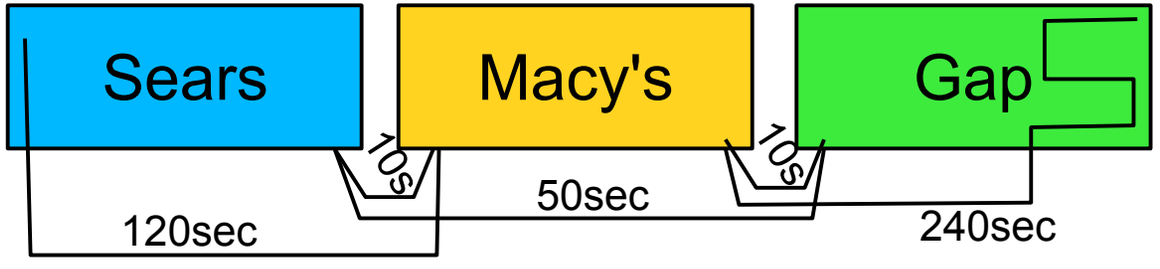
# Some subtleties

- Minimum travel times are now intervals and non Δ=
  - Min TT(Sears → Macy's) = 10 sec
  - Min TT(Macy's → Gap) = 10 sec
  - → Min TT(Sears → Gap) <= 20 sec




- The triangle inequality for points fails for intervals of time/space
  - TT(Sears → Macy's) = [10 sec – 120 sec]
  - TT(Macy's → Gap) = [10 sec – 240 sec]
  - → Min TT(Search → Gap) <= 360 sec (max of Min TT intervals)
- Uncertainty of space modeled by non-Δ= times

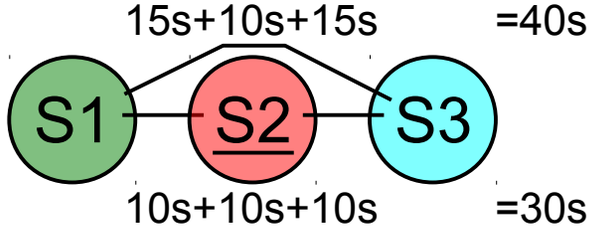


Management Analytics

# Some subtleties

- The absence of evidence **is** evidence of absence (sort of)
  - Covert travel time (CTT)




    - Travel time avoiding detection by (appearances at) sensors
  - Ultimately limits the sequences of locations
  - Also reflects issues associated with missing records
  - May be usable for eliminating unreliable records of various sorts
- Closed spaces and involuntary recording
  - Closed spaces (cannot pass in one (both) direction(s) unsensed)
    - Appearance “inside” requires prior appearance at “perimeter”
    - Appearance “outside” between appearances “inside” without appearance at “perimeter” is inconsistent
    - Closed spaces is a special case of CTT with min CTT higher than the full range of times considered.

Management Analytics

# Some subtleties

- How do I generate TT functions?
  - From related records (lots of them) and other sources
    - E.g., Google Maps and TT estimates augmented with experiments to determine how much less TTmin is
    - E.g., lots of other records – get the 99.999% minimums and investigate inconsistencies per alternative explanations
  - Measure them
    - E.g., for a facility under control they can be measured and updated as facilities change
    - E.g., for a given case, measure what you don't have
  - Calculate them
    - E.g., Person walking DISTANCE (2.3 mi) at SPEED (2.5 mph)
    - E.g., Bus in Delhi at 6PM on Tuesday (1.2 mph)

Management Analytics

# Just for “physical” world?

- TT applies to computer events as well
  - TT(San Diego, San Francisco) > 0.002 sec (speed of light)
  - TT(XXX router leg to leg) >= 0.015 sec
  - TT(face to face TCP 10 Mbyte file on 100mb Ethernet) >= 1sec
- Time and Space Interval Record Schedule Consistency Analysis for **Atomic Items** without Interactions in **Open/Closed** Spaces with **Bounded Locations**
  - The theory and math don't care whether computer files or people travel through Ethernets or streets.
  - Minimum travel time within computers and networks is a very calculable, repeatable, measurable, and testable
  - Many records are kept regularly
  - Interval results and closed space / CTT results deal well with firewalls, system subversions, missing records, etc.

Management Analytics

# Not just a case-by-case approach

- Feed lots and lots of data into the system
  - All transaction data available
  - Augmented on a case by case basis with discovery
- Perform analysis on lots of records
  - Complexity ~nLog(n) → highly scaleable
- Detect inconsistencies
  - Establish reliability of records as a matter of course
  - Establish criteria for authenticity of records along the way
- Apply to all / many cases, not just specific cases
  - When inconsistency found, resolve the case and the records
  - Improve with time

Management Analytics

# What about John?

- At ATM 6:45 (shown by ATM receipt)
  - 15 minutes from flower store and bookstore
- At Flower store 7:05 (shown by flower store receipt)
  - Flower store 3 minutes from bookstore
- At Mary's apartment at 7:30 (watched news together)
  - 20 minutes from flower store to apartment
- Currently consistent with John at bookstore at 7:00-7:02
- Currently consistent with John at bookstore at 7:08-7:10
  - Traffic cameras did not see John between 7:00 and 7:10
  - Flower store 6 minutes from bookstore covert to cameras
    - John could not have done it? NO!
- → Records unreliable, OR not authentic, OR John didn't do it

Management Analytics

# In context of many cases

- Financial transaction records established as good to within [time interval] with 99.xxxx% probability
- Television news start times established as good to within …
- Traffic camera detection of suspects established as good to within …
- Police time records established as good to within …
- Consistency of records analysis method established as good to within …
- Draw reasonable conclusions from records
- Rapidly determine to bring case against John (or not)

Management Analytics

# What's next?

**Time and Space Interval Record Schedule Consistency Analysis for Limited non-Atomic Items with/out Interactions in Open/Closed/Covert Spaces with Non-Stationary Locations**

- With interactions implies changes based on interactions
  - e.g., additional location(t) records by effects of interactions
    - Jane saw Jill with Jack in the park at about 3PM
    - The ordering of events associated with causal chains
      - e.g., Windows records this before that, etc.
- Limited non-atomic objects
  - e.g., Program uses a 25 libraries, but not all of them all the time.
  - e.g., Susan and I share a phone and credit cards while in Delhi
- Note that as we add more of these improvements, all the old methods remain valid. We just get tighter bounds on times and better detection of inconsistency.

Management Analytics

# Thank You

**http://all.net/ - fc at all.net**