International Symposium on High Confidence Software, Beijing, Dec 2011

The multi-facets of building dependable applications over connected physical objects

S.C. Cheung
Director of RFID Center
Department of Computer Science and Engineering
The Hong Kong University of Science and Technology
http://www.cse.ust.hk/~scc

Due to my own experience, I will draw examples from RFID applications for illustration.

Yet, the concepts and issues discussed should be generalizable to the physical objects connected using other wireless technologies, such as wifi, zigbee, bluetooth and so on.

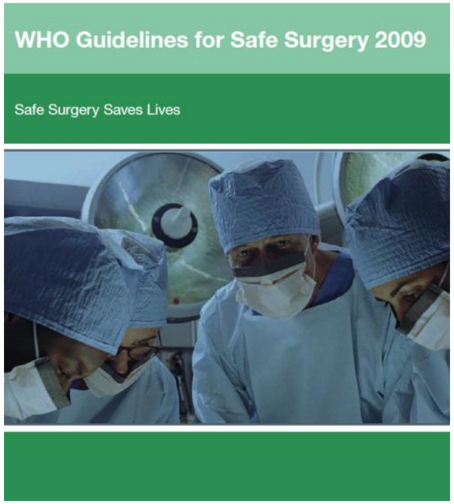
Where are the needs for connected physical objects?

To meet the demands on high confidence applications

WHO Guidelines for Safe Surgery 2009

- Objective 7: The team will prevent inadvertent retention of instruments or sponges in surgical wounds.
  - Need to systematically track/count the sponges used in an operation.

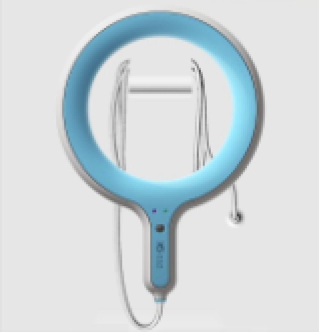
World Health Organization | Patient Safety



RFID-enabled sponges

- Real-time tracking and counting sponges used in an operation

RFID-enabled sponge



Connected physical objects (CPO)

- Stay connected.
  - Multiple networking technologies: Wifi, Zigbee, Bluetooth, RFID, …

An RFID gate that consists of four antennas

Loading bay Storage bay

- Identify themselves actively.
- Feature tight combination of computational and physical elements.
- Tight integration of computation and physical processes.

Connected physical objects (CPO)

Call for interdisciplinary efforts:

- Algorithm
- Database
- Pervasive computing
- Logistics management
- …
- **Software engineering**

Question 1

Are there any fundamental differences between CPO systems and conventional software systems?

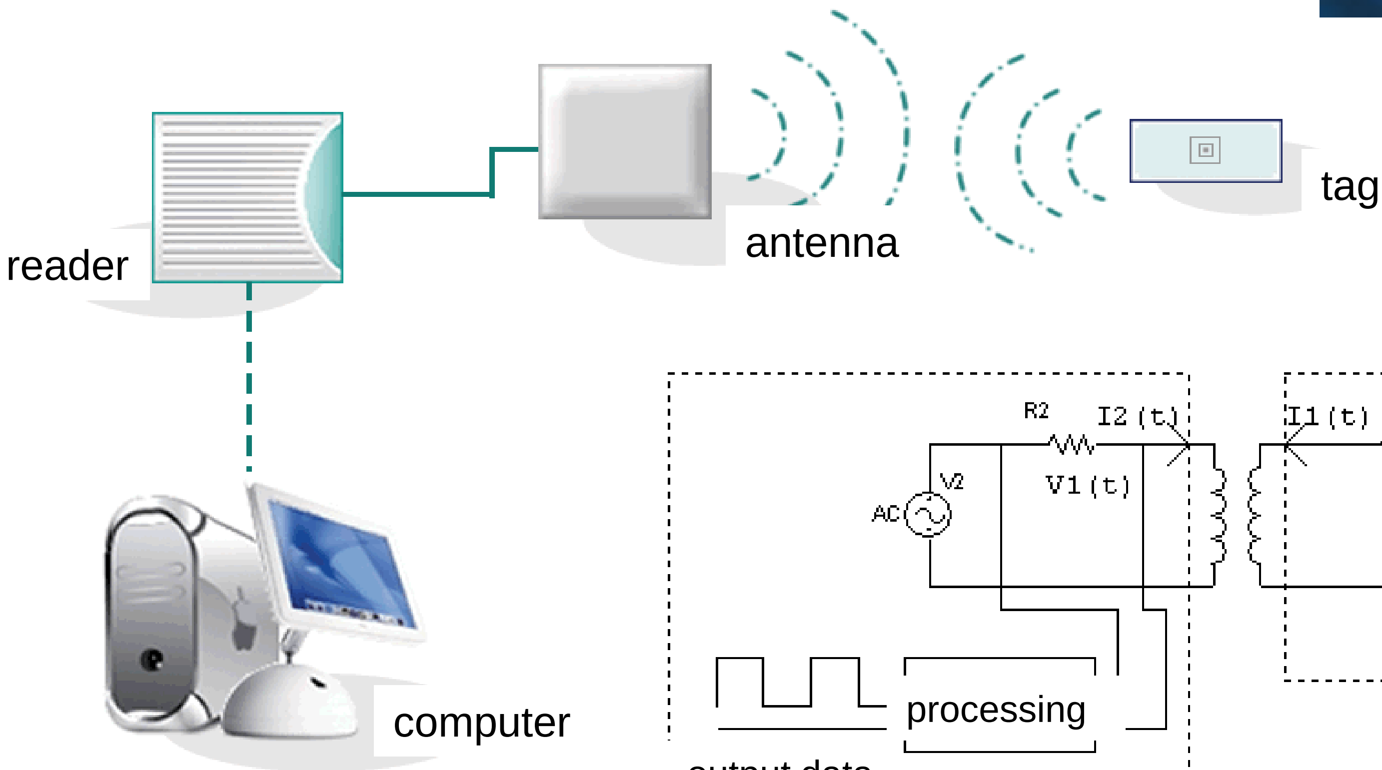
Question 2

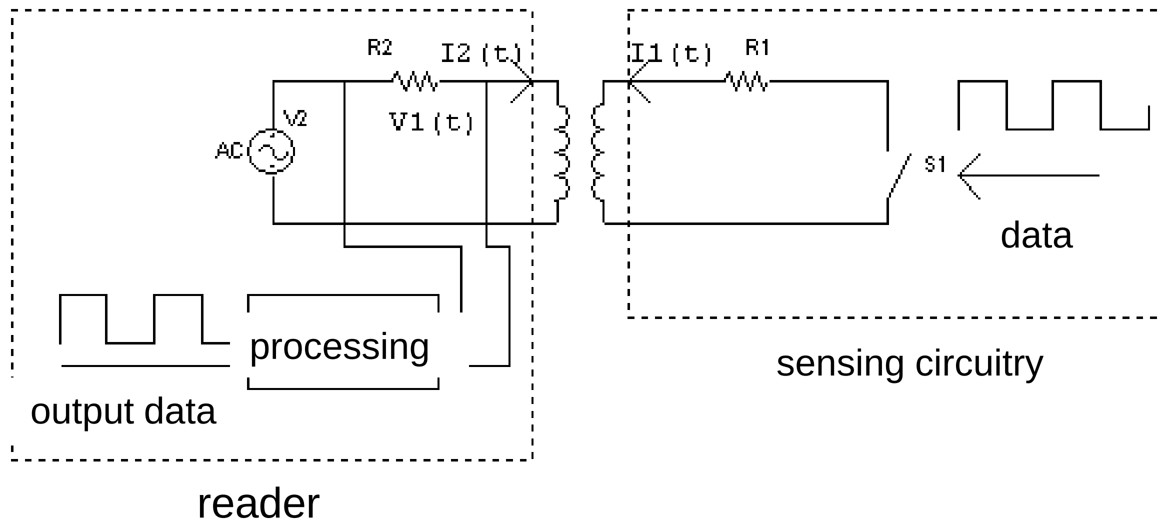
In which way CPO systems induce new challenges to high confidence software?

Observation

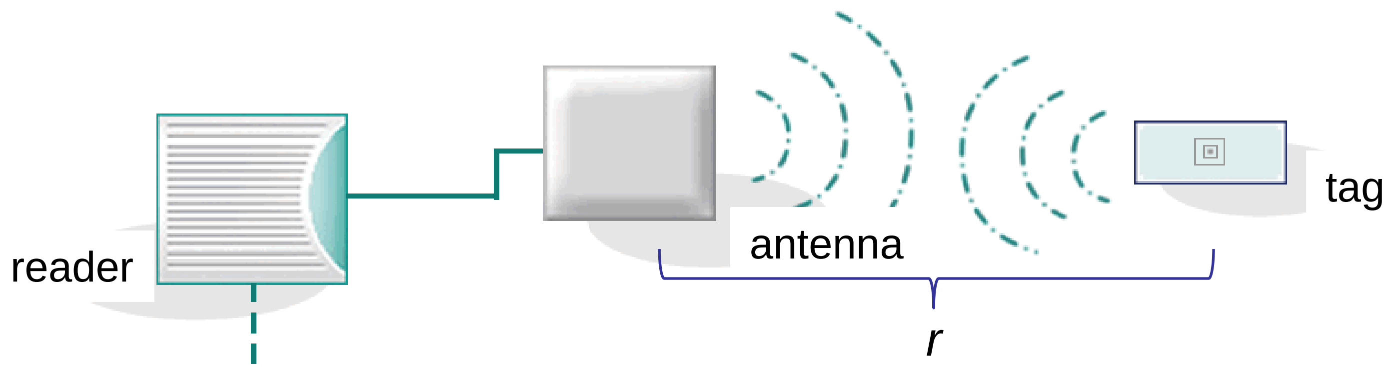
By being physical in nature, CPOs are subject to physical laws and able to sense or interact with their physical environment

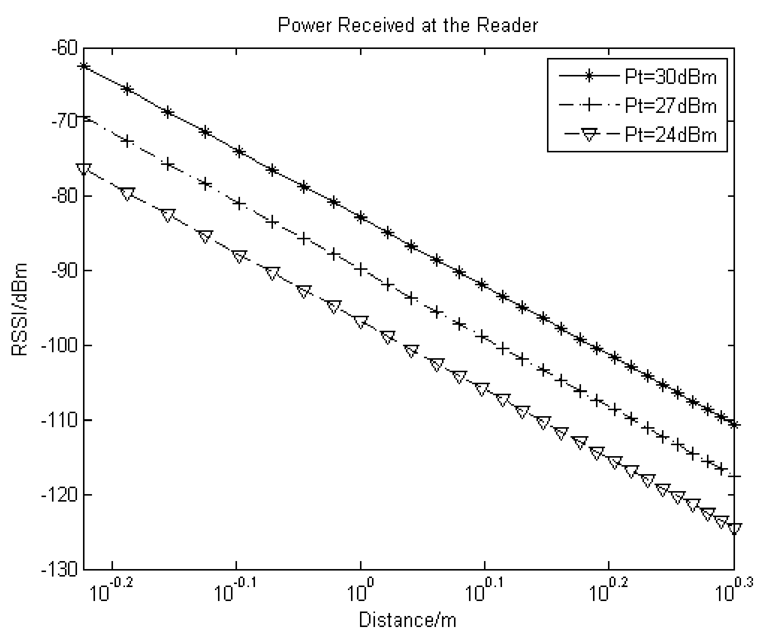
Physical law





Physical law





transmit power

wavelength

$$P_{RX} = P_{TX} G_{TX} G_{RX} \left(\frac{\lambda}{4 \pi r} \right)^2$$

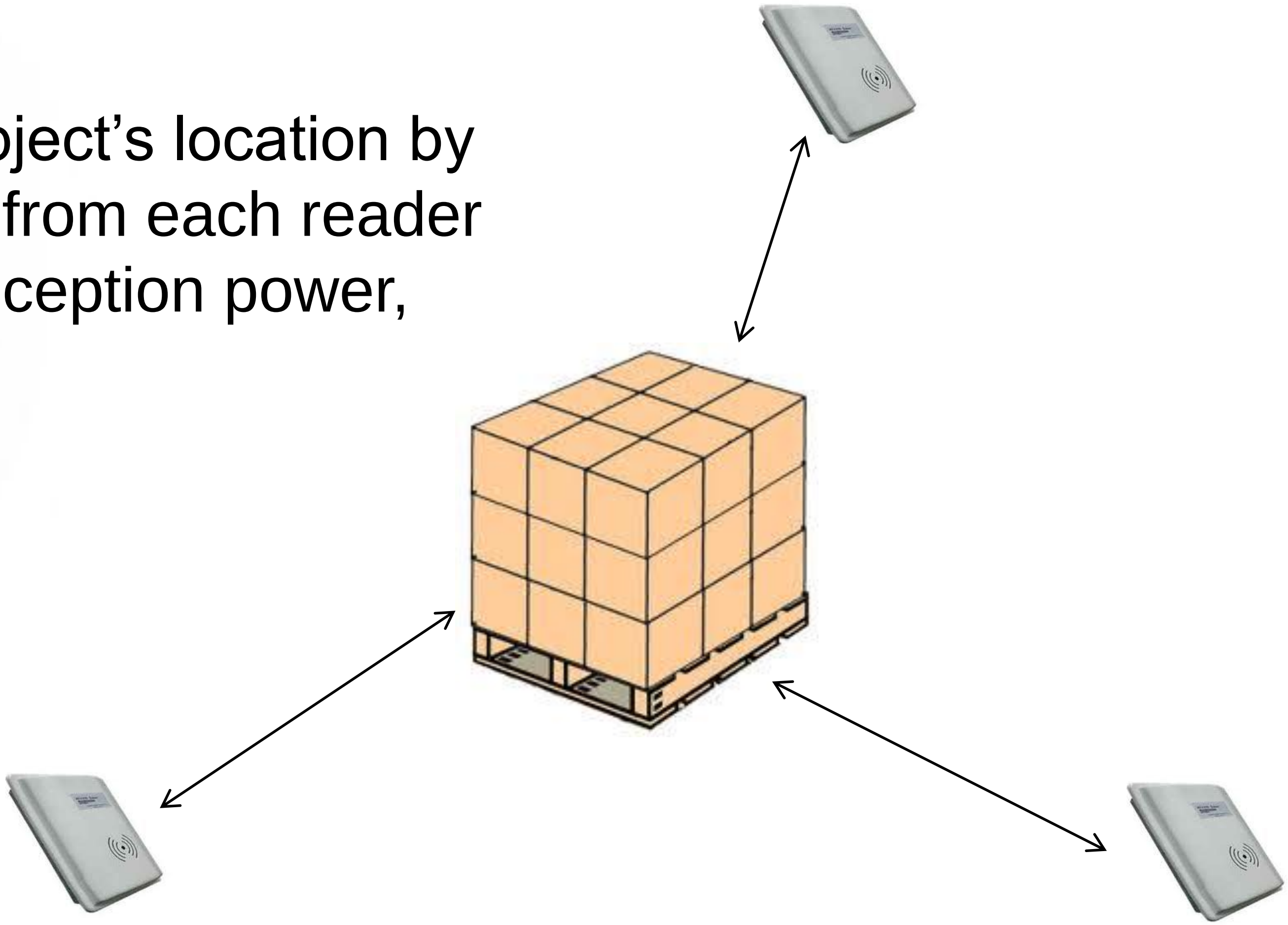
received power

gain of transmit and receive antennas

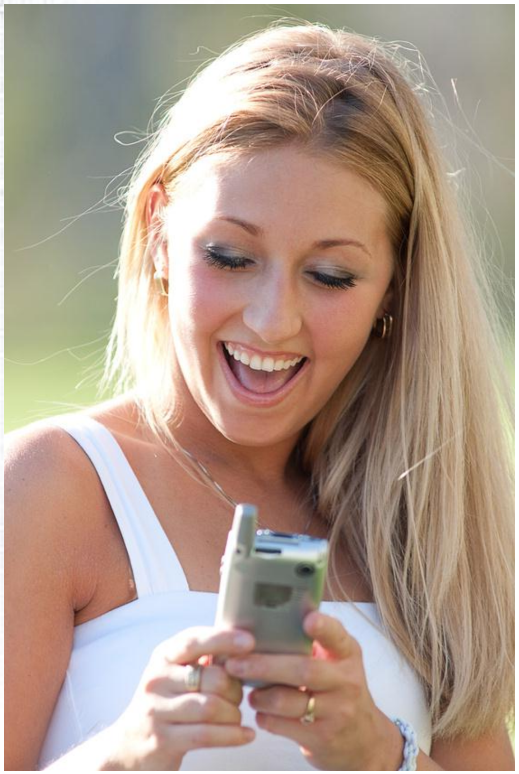
separation distance

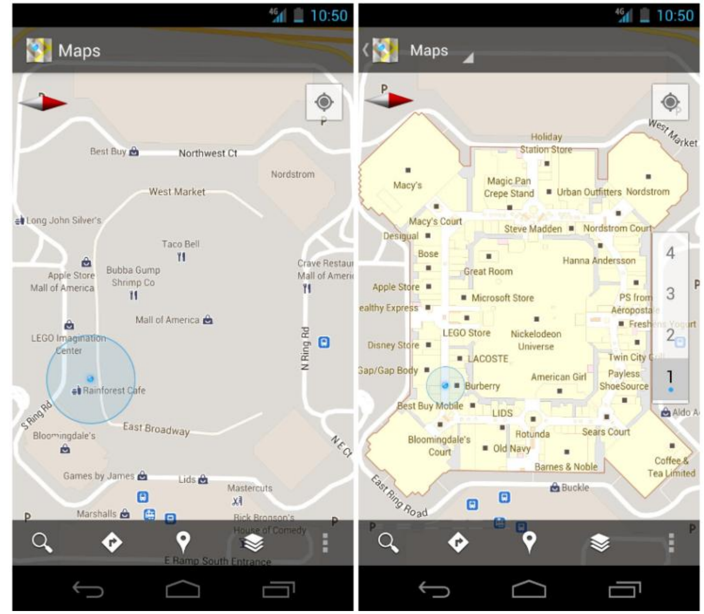
Object tracking

Compute object's location by its distance from each reader based on reception power, e.g., RSSI.



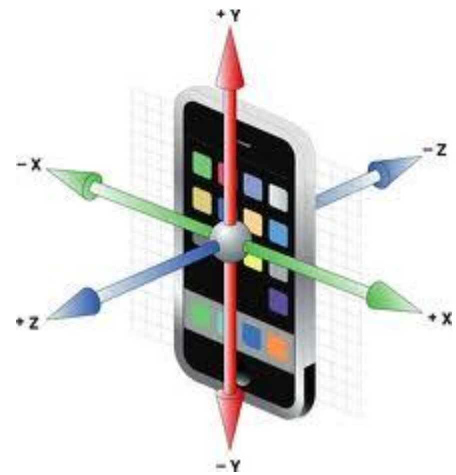
Finding directions

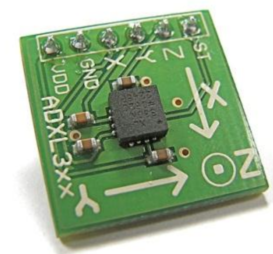




Sensing environment

accelerometer

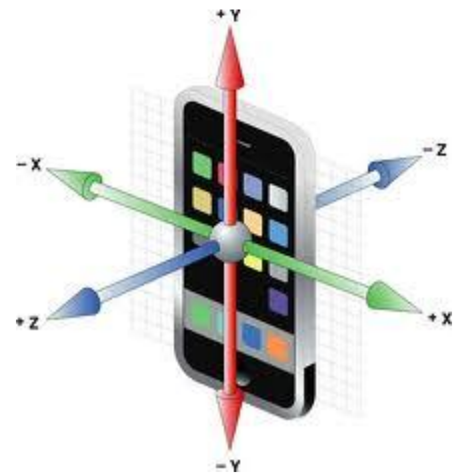


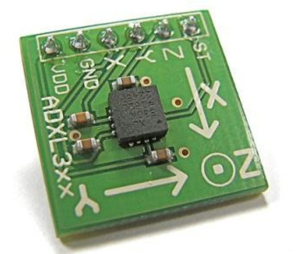


About 35 percent of people over age 65 fall in their homes at least once each year. That figure increases to 50 percent for those with age 75 and over. -- Havard Health Letters 2009

Adaptive actions

accelerometer

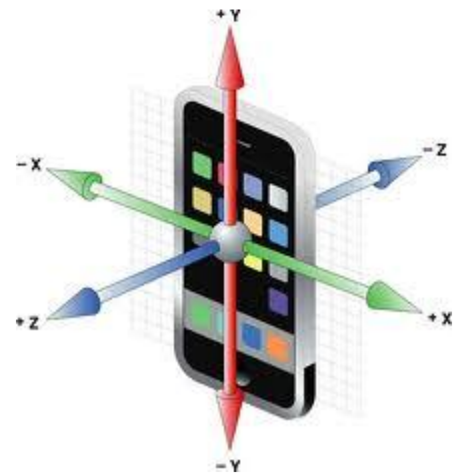


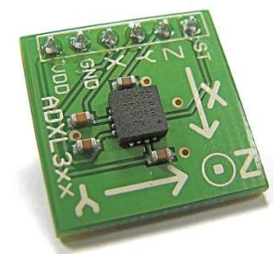


- Raise an alarm
- Call her relatives
- Call her neighbors
- Call emergency services
- Contact her nearest hospital
- Call her family doctor
- Record the frequency

Adaptive actions

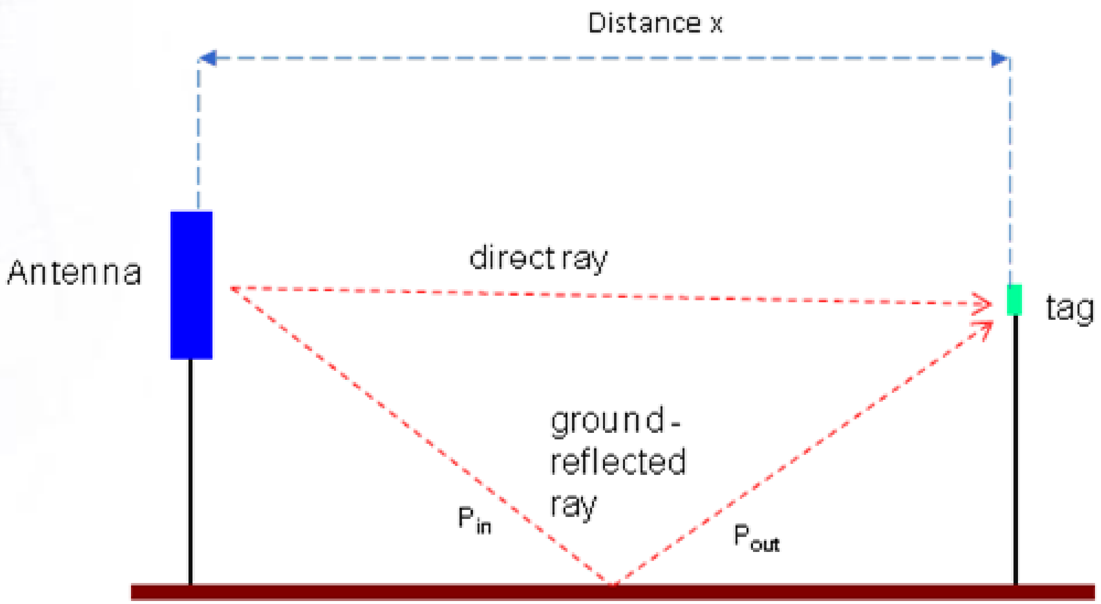
accelerometer





- Raise an alarm
- Call her relatives
- Call her neighbors
- Call emergency services
- Contact her nearest hospital
- Call her family doctor
- Record the frequency

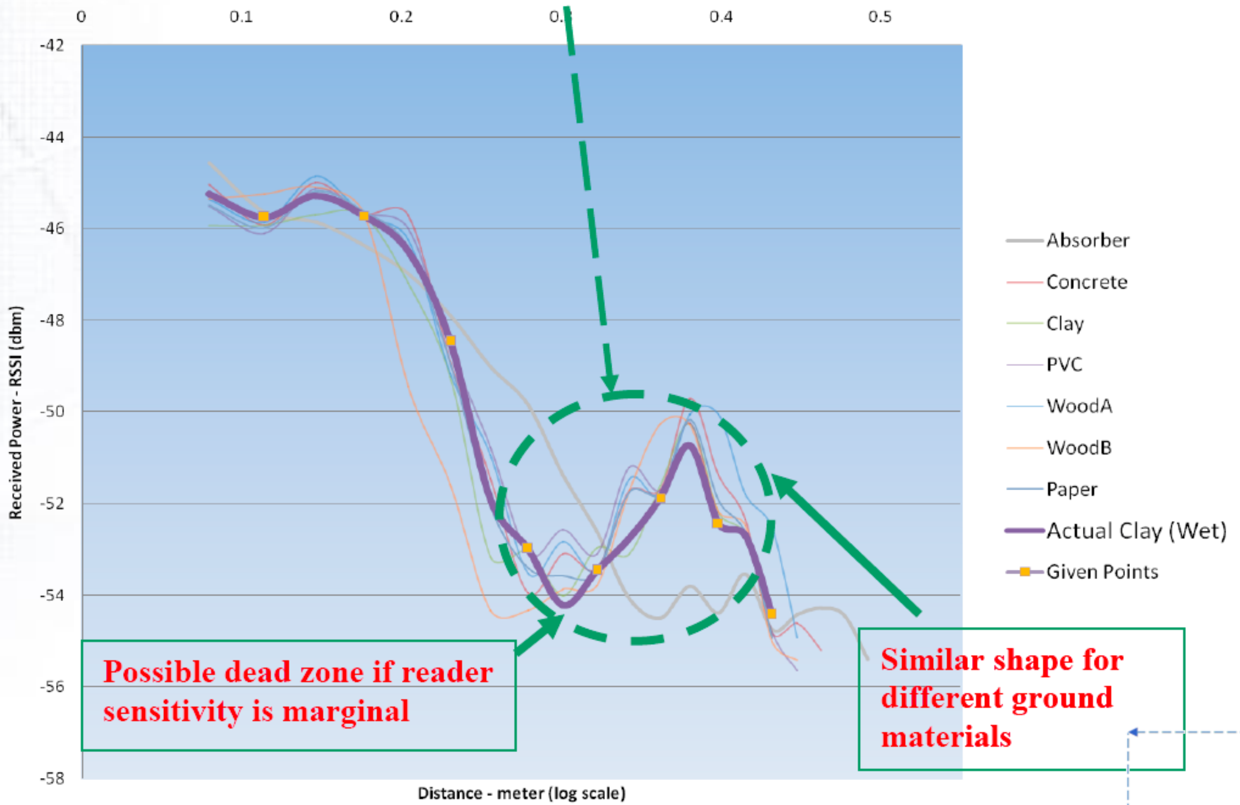
Subject to physical laws

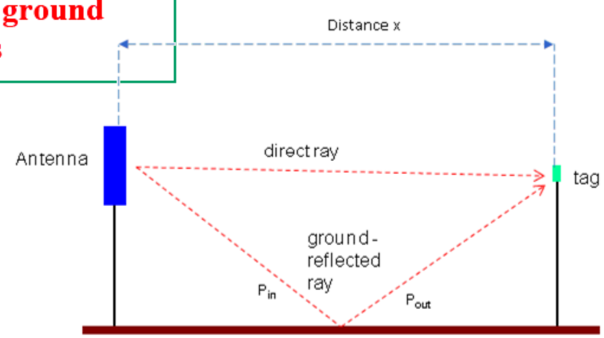


Multi-path Effect

Subject to physical laws

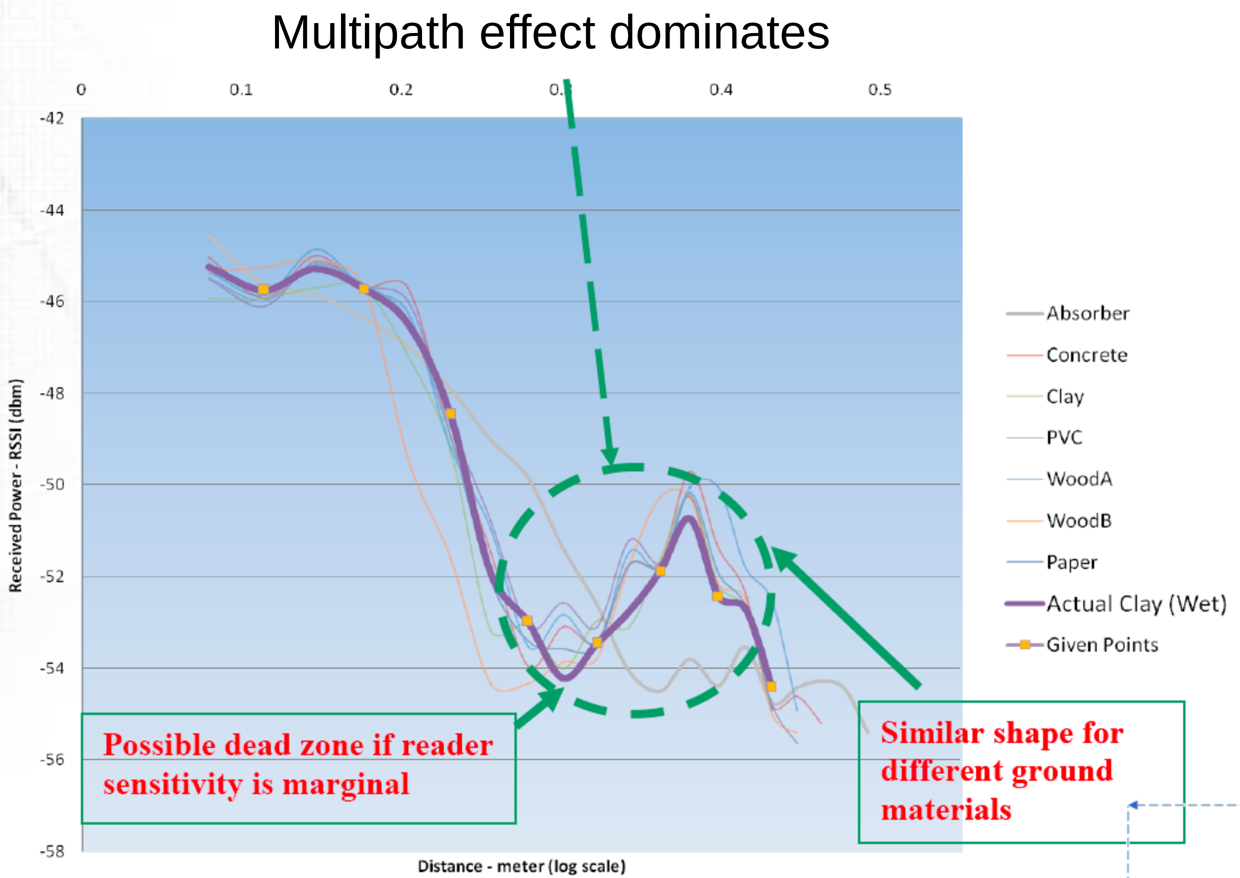
Multipath effect dominates

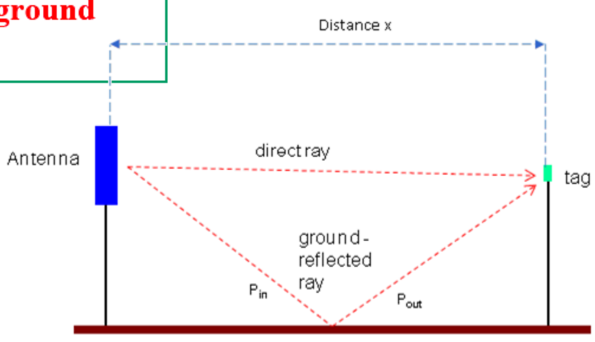




- Induces noise to data signals.

Subject to physical laws





- Leads to inaccuracy in object tracking.
- Jumping on the map when using GPS

Induces engineering problem

Induces engineering problem

Is the problem common?

- For RFID data, the observed read rate (i.e., the percentage of tags in a reader's vicinity that are actually reported) in real-life deployment may drop down to the range of 60-70% [1]

[1] Jeffery, S.R., Garofalakis, M., and Frankin, M.J. Adaptive Cleaning for RFID Data Streams. In *Proceedings of the 32nd International Conference on Very Large Data Bases*, Seoul, Korea, September 2006, 163-174.

New challenges

- Data level
  - Ubiquitous sensing noises are not under full control.
  - Natural constraints on physical objects cannot be fully and precisely modeled.
- Logic level
  - Hardly able to consider adequately the real-life complexity of physical objects that are evolvable.
  - Unanticipated exceptions occur.

Complexities arise from CPO

- Data level
  - Ubiquitous sensing noises are not under full control.
  - Natural constraints on physical objects cannot be modeled completely and precisely.
- Logic level
  - Hardly able to consider adequately the real-life complexity of physical objects that are evolvable.
  - Unanticipated exceptions occur.

Complexities arise from CPO

- Data level
  - Ubiquitous sensing noises are not under full control.
  - Natural constraints on physical objects cannot be modeled completely and precisely.

➔ Sensing defects

- Logic level
  - Hardly able to consider adequately the real-life complexity of physical objects that are evolvable and noisy.
  - Unanticipated exceptions occur.

Complexities arise from CPO

- Data level
  - Ubiquitous sensing noises are not under full control.
  - Natural constraints on physical objects cannot be modeled completely and precisely.

➔ Sensing defects

- Logic level
  - Hardly able to consider adequately the real-life complexity of physical objects that are volatile and noisy.
  - Unanticipated exceptions occur.

➔ Adaptation defects

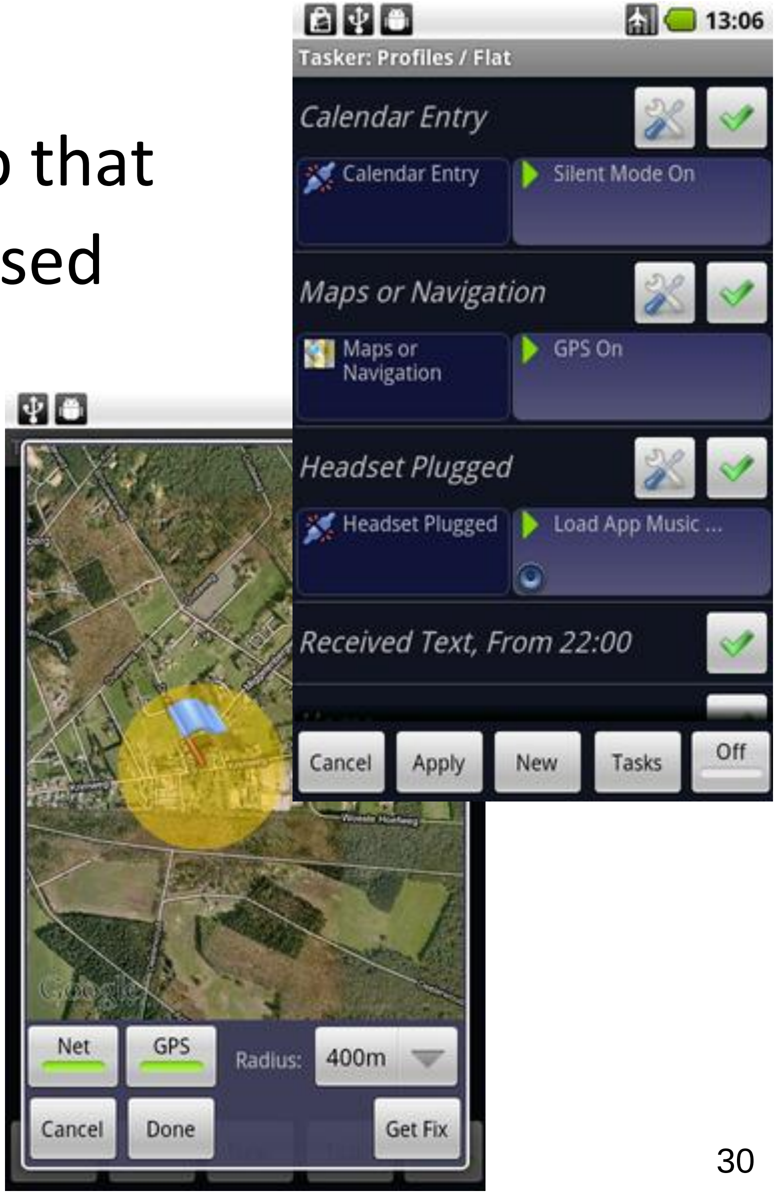
Sensing and adaptation defects are common to CPO systems.

Sensing defects

- Tasker is a popular Android app that performs user-defined tasks based on its environment changes (or situations).

An early bug (in v.20b):

- Due to a sensing error on wifi.
- Kept trying to establish connection.
- Drained battery power quickly.

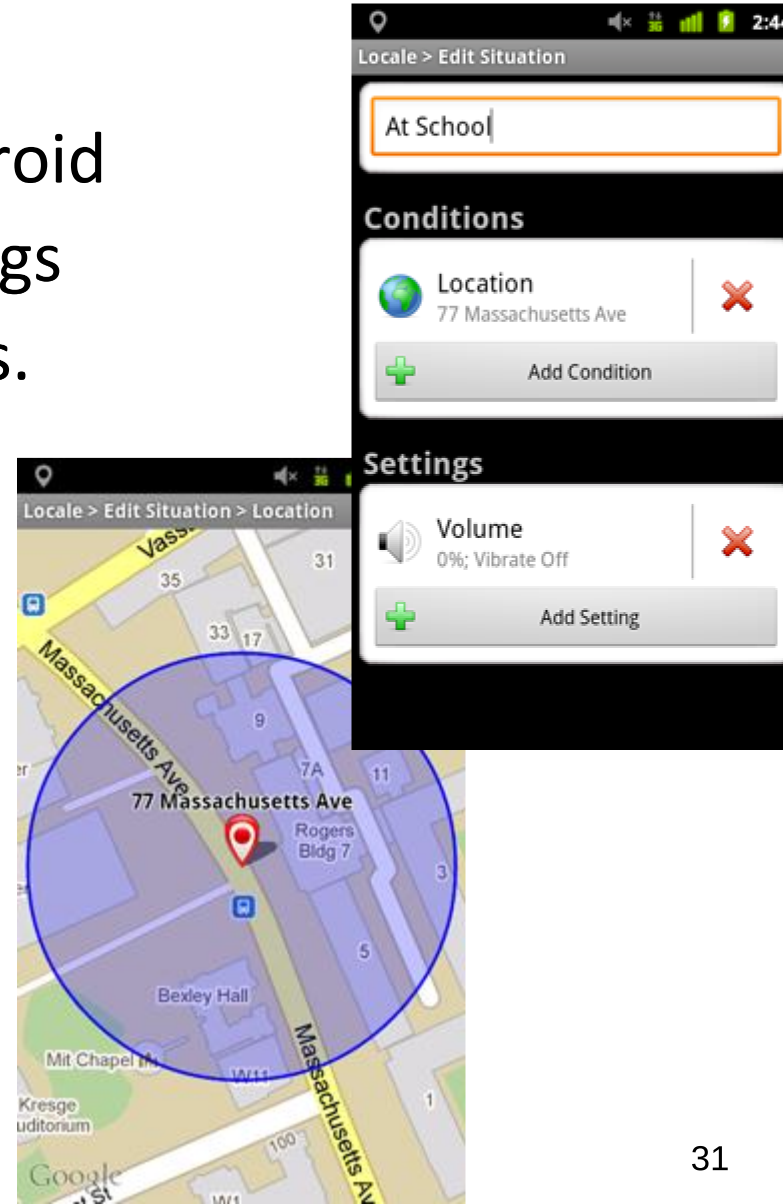


Adaptation defects

- Locale is an award-winning Android app that changes phone's settings based on pre-defined conditions.

An early bug (in v1.0):

- Mistake in condition evaluation.
- May bounced rapidly between adapted settings and default settings occasionally.
- Drained most of the phone processing power.
- Failed to respond to user inputs.

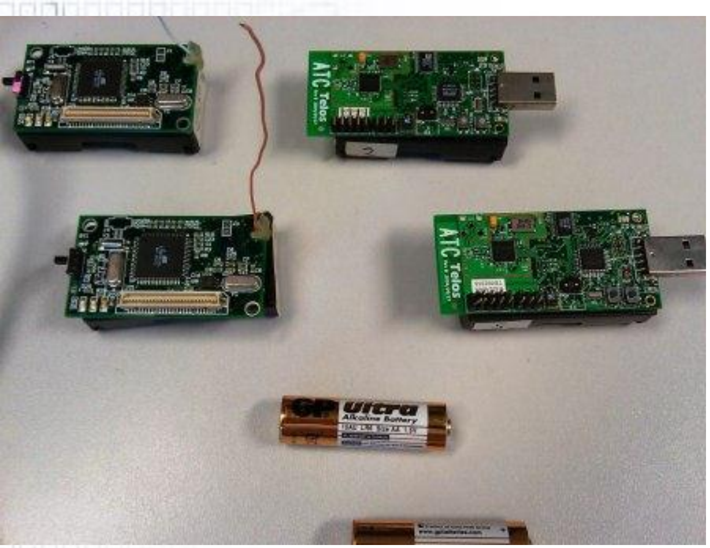


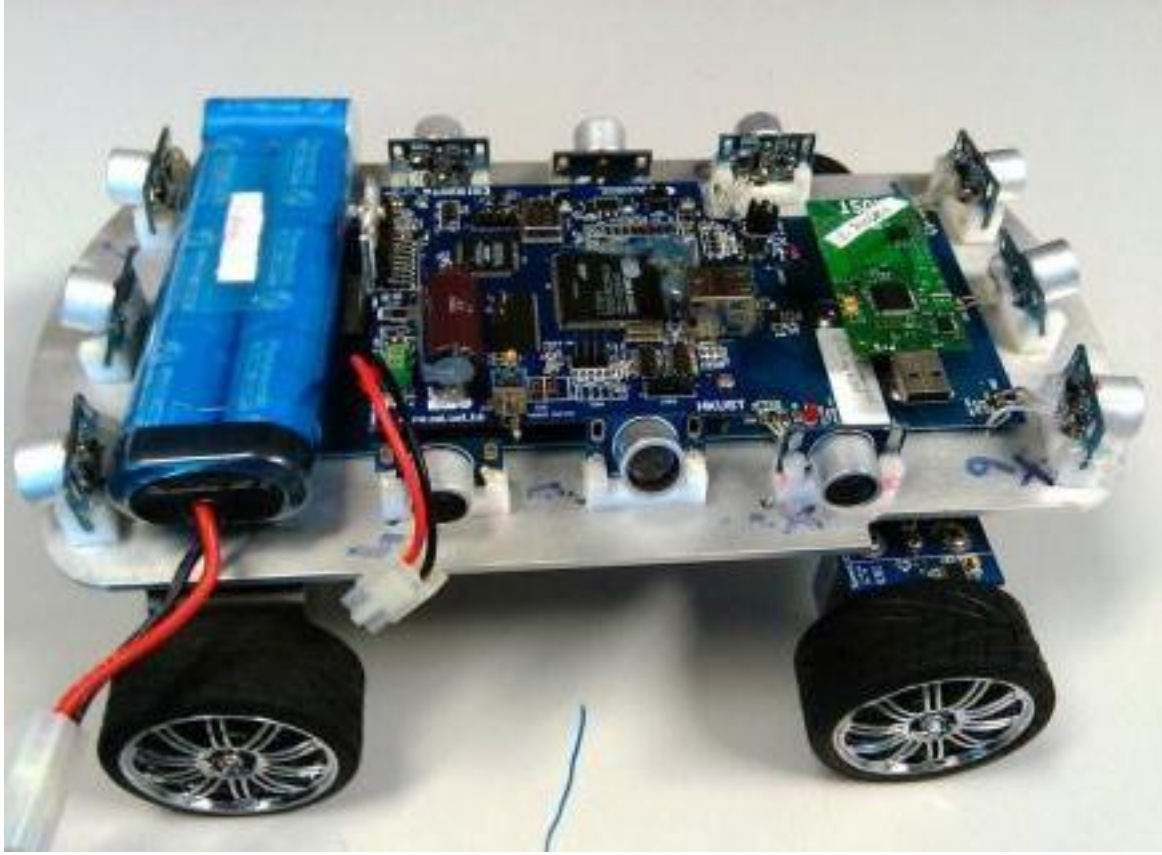
Defect detection and management

Defect detection

- In most cases, we don’t have the ground truth.
- We are not able to tell if a sensed datum (or context) is correct.
  - The meaning of “correct” is unclear.
  - We often live with some errors.
- CPO are subject to physical rules.
- Extract constraints that cover context changes from these rules.
- Catch context inconsistency upon rule violation.

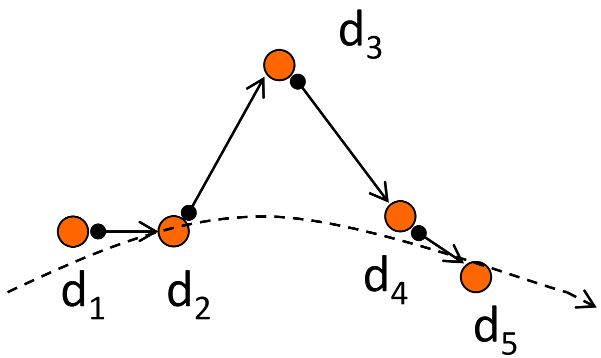
Illustrative Example

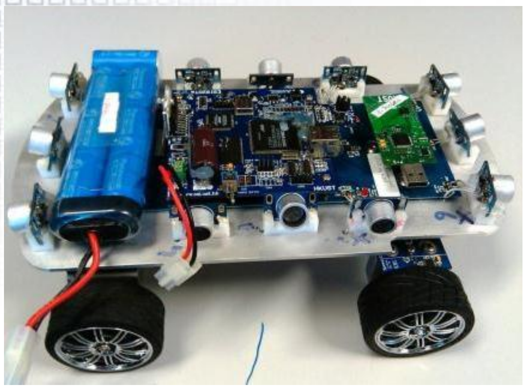




Illustrative Example

It is hard to tell which of the five contexts is/are incorrect unless we know the ground truth.



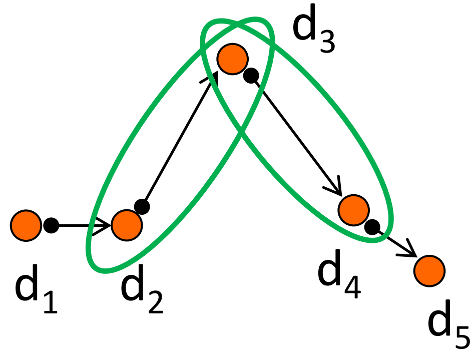


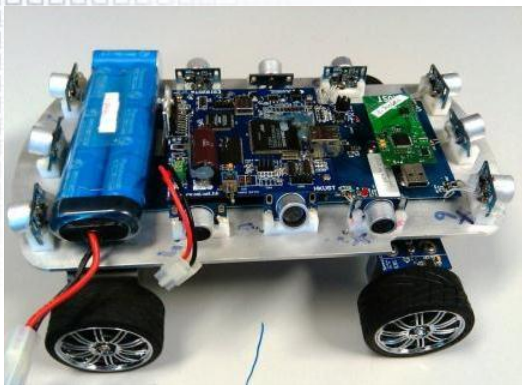
The minicar reports that it has visited five locations

Illustrative Example

Constraint: The walking speed estimated from location changes should be less than 150% · v

The average speed obtained from the sensed contexts





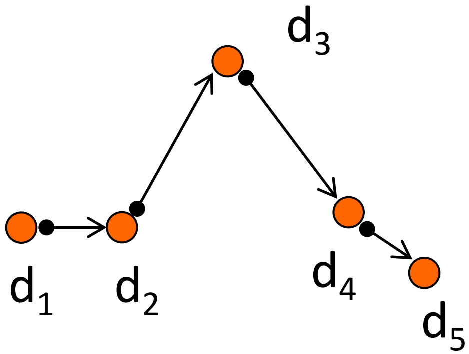
Context changes
(d_2, d_3), (d_3, d_4)
violate the constraint

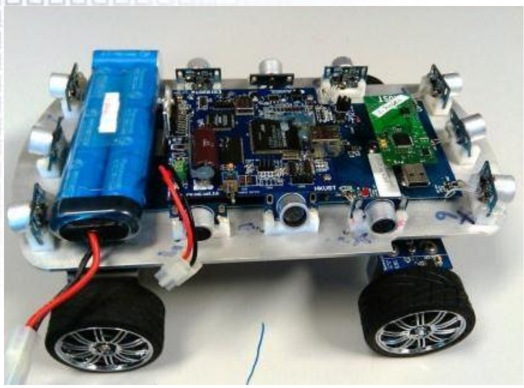
Two context inconsistencies found

Illustrative Example

Constraint: The walking speed estimated from location changes should be less than 150% · v

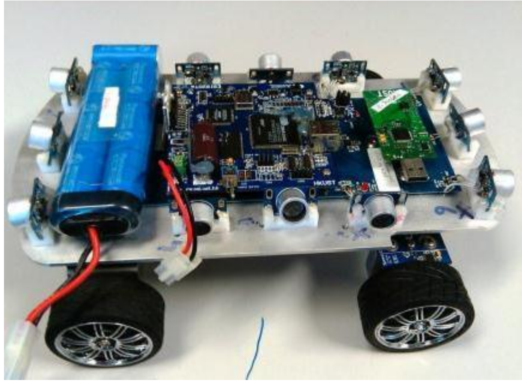
The average speed obtained from the sensed contexts





Useful constraints are those that can affect the adaptive behavior of connected physical objects

Research questions



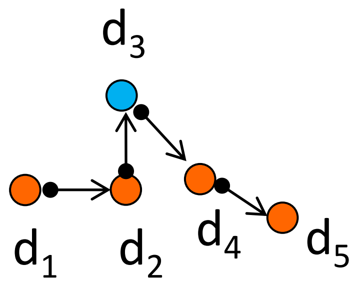
- How to identify useful constraints?
- How to formulate constraints?
- How to model contexts?
- Do we have a time window on the contexts?
- How to evaluate constraints in view of new contexts?
- May multiple context changes be evaluated in different order?
- Can the evaluation be incrementally done?
- What if constraints are changed dynamically?
- May context inconsistencies be resolved?

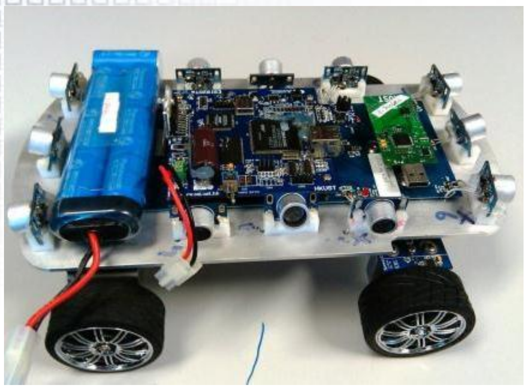
[Xu et al. ESEC/FSE 2005, ICSE 2006, SEAMS 2007, ICDCS 2008, TOSEM 2010]

What if adaptation defects?

Likely induce transient bugs

Hard to catch without a test oracle

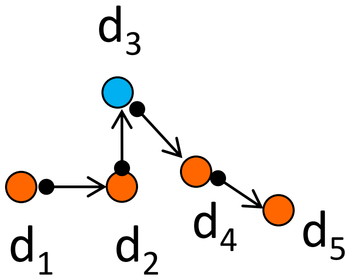




The minicar reports that it
has visited five locations

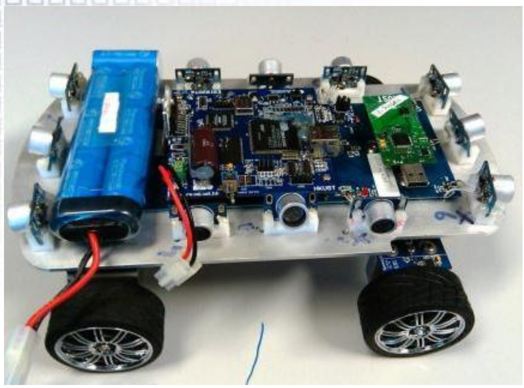
Interval-based analysis

- Partition context streams into intervals
- Cluster these intervals based on their physical and logical features
- Intervals that appear as outliners likely contain transient bugs



$d_1d_2d_3d_4d_5d_6\ldots\ldots\ldots.d_nd_{n+1}$

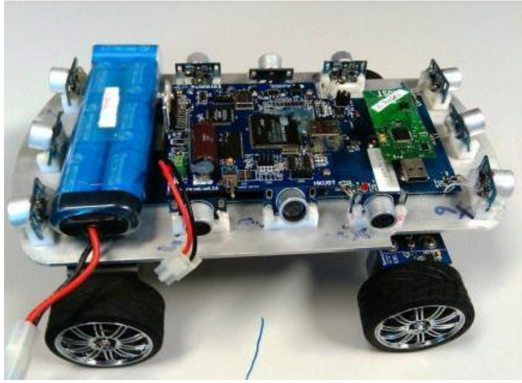
$|d_1d_2d_3d_4d_5|\ |d_6\ldots d_{12}|\ \ldots\ |\ldots d_nd_{n+1}|$



The minicar reports that it has visited five locations

[Ngai et al. Ad Hoc Networks 2010; Zhou et al. ICDCS 2010]

Fault detection of adaptation logic



- How to model adaptation logic at design stage?
- Can we detect adaptation logic at design stage?
- What are adaptation design faults?
- What are the causes of adaptation design faults?
- Do these faults have patterns?
- Can we statically check adaptation design faults?
- Can we generate test cases for CPO systems from their design?
- Do we have new test coverage criteria for CPO systems?

[Lai et al. FSE 2008; Sama et al. TSE 2010, JSS 2010; Wang et al. ICSE 2007]

Conclusion

- Physical objects are no longer static, passive data; they are ***intelligent agents*** that interact with each other and our traditional software artifacts
  - The ***data*** of physical things (e.g., states, properties) are from everywhere, with noisy nature
  - The ***logics*** of physical things (e.g., how to interact with each other) are distributed everywhere
- There is an inevitable impact on application development, software testing, and fault localization
  - E.g., application code is distributed anywhere, composed at runtime, and subject to change

More research challenges

[Ma, Baresi, Ghezzi et al. ESEC/FSE 2011]

- How to update adaptive components dynamically?
- How to model such requirements?
- What is the role of middleware?
- Can middleware be dynamically updated as well?
- What if exceptions occur?
- Can new programming methodology help?
- What is a test case?
- How to define test oracles in a cyber-physical world?

More research challenges

- How to update adaptive components dynamically?
- How to model such requirements?
- What is the role of middleware?
- Can middleware be dynamically updated as well?
- What if exceptions occur?
- Can new programming methodology help?
- What is a test case?
- How to define test oracles in a cyber-physical world?

The list is by no means exhaustive

More research challenges

- How to update adaptive components dynamically?
- How to model such requirements?
- What is the role of middleware?
- Can middleware be dynamically updated as well?
- What if exceptions occur?
- Can aspect-oriented development help?
- What is a test case?
- How to define test oracles in a cyber-physical world?

Solutions to these challenges likely involve knowledge beyond software engineering and even computer science

Preliminary work at HKUST

video

Our RFID Test Labs

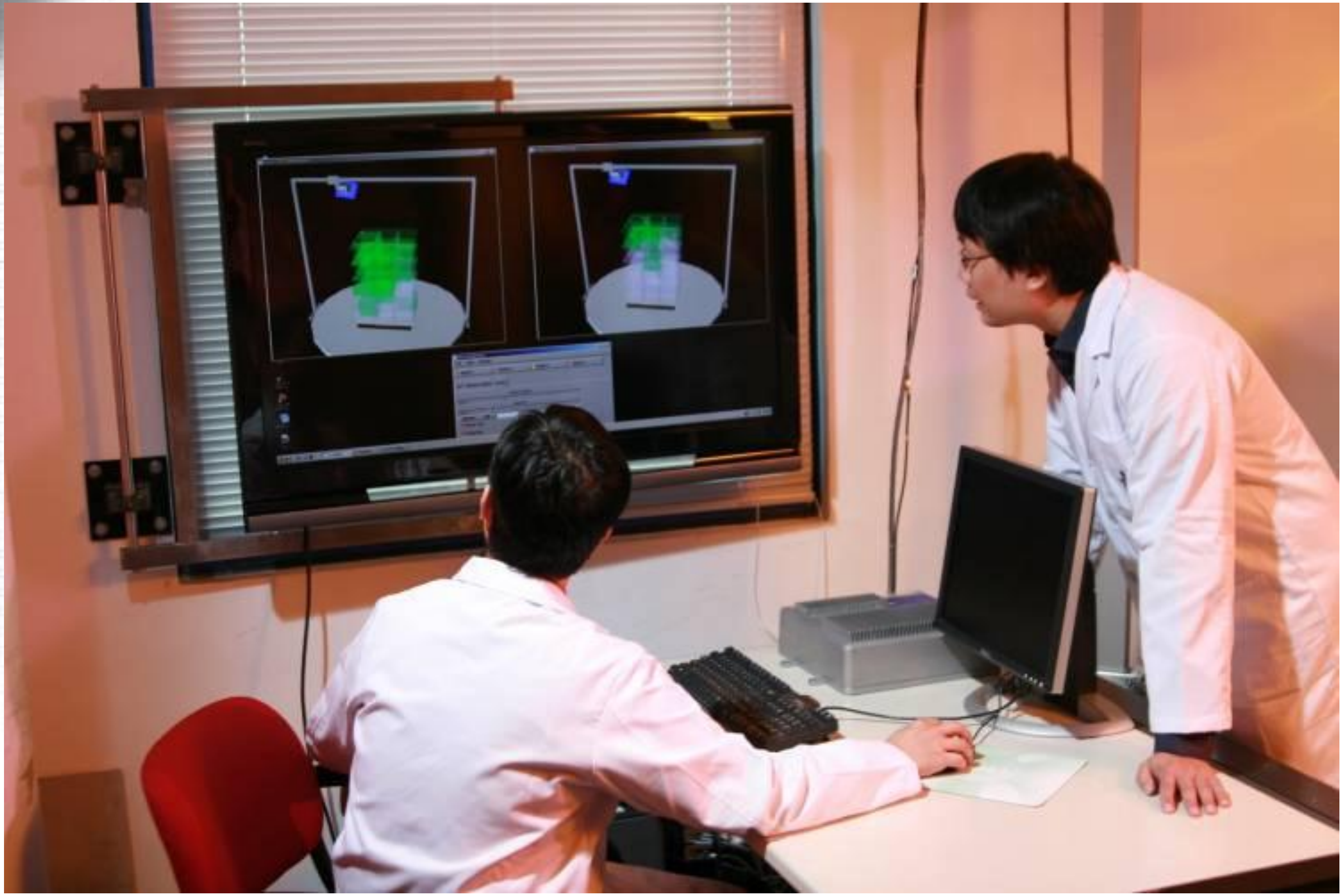
- Clear Water Bay (HKUST) Lab
- Nansha (HKUST) Lab
- Founding member of the Global RF Lab Alliance (www.grfla.org) with other 7 international research labs

RFID Test Lab (Clear Water Bay)

RFID Test Lab (Clear Water Bay)

RFID Test Lab (Clear Water Bay)

RFID Test Lab (Clear Water Bay)



Test Lab (Nansha)

Thank you

Abstract

The advances of wireless and device technologies have led to seamless integration of physical objects that are connected using wireless networks. Examples of physical objects include sensors, Radio Frequency Identification tags, and transceivers that obtain geographical location information. Physical computing systems are applications that built upon these physical objects which intercepts real-time information of their environment. The information can be used by physical computing systems to adapt their behavior. However, construction of dependable physical computing systems poses new challenges such as sensing ranges, power-awareness, computational models, and object mobility. These challenges bring interesting research opportunities to program analysis, testing, modeling and verification. In this talk, we will introduce several emerging physical computing systems, review major challenges and discuss preliminary efforts.

Biography

Dr. Shing-chi (S.C.) Cheung received his BEng degree in Electrical Engineering from The University of Hong Kong, and his Ph.D degree in Computing from the Imperial College London. In 1994 he joined The Hong Kong University of Science and Technology (HKUST) where he is a full professor in the Department of Computer Science and Engineering, and the director of RFID Center. He participates actively in the program and organizing committees of major international conferences on software engineering. He was an editorial board member of the IEEE Transactions on Software Engineering (TSE, 2006-9).

He is interested in studying software engineering issues related to program analysis, testing and debugging, services computing, cyber-physical systems, internet of things, and mining software repository.