Safety and Privacy Matter for Human Cyber-Physical Systems
How to Build in Dependable Autonomy

Zhi Jin

Key Lab of High Confidence Software Technologies (MoE)
Peking University
zhijin@pku.edu.cn

高可信软件技术教育部重点实验室

DSA 2020，Xian, China, 2020.11.28

北京大学

Outline

Motivation

Survey

Thinking about Design

Conclusion

Cyber-Physical Systems and Human in the Loop

- HCPS are engineered systems
  - that are built from, and depend upon the seamless integration of computation and physical components and human in the loop
- HCPS technologies are
  - transforming the way people interact with engineered systems
- HCPS drive
  - innovation and competition in a range of application domains including
  - agriculture, aeronautics, building design, civil infrastructure, energy, environmental quality, healthcare and personalized medicine, manufacturing, and transportation

Human Cyber-Physical Systems: Hot Topic

Smart Anything Everywhere

- Since 2014
- Advanced micro-electronic components and smart system integration
- Organic and large area electronics
- Cyber-physical and embedded systems
- Customized low energy computing powering CPS and the Internet of Things

DARPA Robotics Challenge

- Since 2015
- develop human-supervised ground robots capable of executing complex tasks in dangerous, degraded, human-engineered environments

Human Cyber-Physical Systems: Hot Topic

The Smart and Autonomous Systems Program: Since 2015

- Invests $10 million in smart, human-centered service systems
- Spur innovation for smart, manufacturing and infrastructure

The Cyber-Physical Systems and Smart and Connected Communities Program: Since 2019

- Small: up to $500,000. emerging new and innovative ideas that may have high impact on the field of CPS.
- Medium: $500,001 to $1,200,000. multi-disciplinary projects that accomplish clear goals requiring integrated perspectives spanning the disciplines.
- Frontier: address clearly identified critical CPS challenges that cannot be achieved by a set of smaller projects. look to push the boundaries of CPS well beyond today's systems and capabilities. $1,200,001 to $7,000,000.

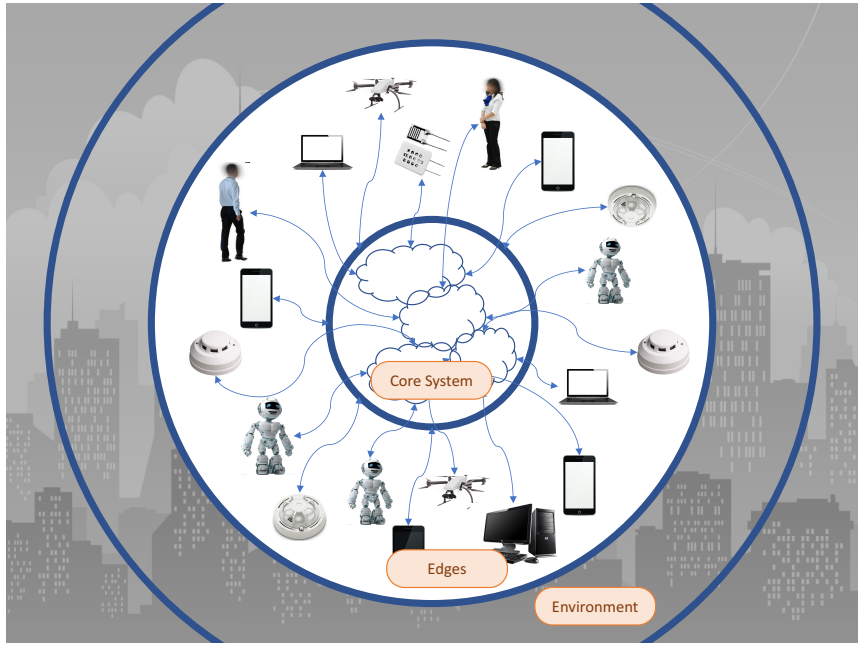
National Science Foundation
WHERE DISCOVERIES BEGIN

Application Examples

- Increasing number of critical applications in dependable sectors
  - Smart transportation
  - Large-scale critical infrastructures
  - Intelligent defense systems
  - Smart health care
  - aviation
  - ……
    - ……

Featured Requirements

- HCPSs are featured by:
  - Spatial distribution
  - Direct interactions with environment
  - Autonomous information sharing and task coordination



Spatial Distribution

- Computing goal
  - be achieved by a set of diverse computing tasks
- Computing tasks
  - are assigned on different computing units
- Computing units
  - are spatially distributed but networked
  - communicate and coordinate their actions for achieving a common goal

Direct Interactions

- Between humans and their devices by
  - providing services of notification, personalized use, context sensing, and augmented reality
- Between humans
  - supporting negotiation, joint planning, collaborative task execution, learning, and direct democracy
- Between devices and software services
  - data analytics, service compositions, and micro-transactions
- Between system and environments
  - sensing, and actuation

Autonomy, Self-Adaptation

- Be capable of robust, long-term autonomy requiring
  - minimal or no human operator intervention
  - in the face of
    - Uncertain, unanticipated, and dynamically changing situations
- Combine
  - perception, cognition, communication, and actuation

  to operate in physical world

New Paradigm and New Challenges

- Environmental sensing and human behavior inclusion: lead to
  - the design of novel safe methods for including humans in the data analysis, processing and decision-making loop
  - the design of distributed adaptation to human behaviors in different contexts
  - the design of the secure and sensitive-aware edge data analytics respecting privacy
- The coexistence of trusted or partially trusted nodes with malicious ones: lead to
  - the design of novel security mechanism, that may be distributed compared with that in traditional centralized computing
- Large-scale nodes coordination: lead to
  - correct and efficient tradeoff between of computing and communication with taking care of other issues, like energy issue
  - Combination of scalability with security, safety and privacy in massive overlays

Outline

Motivation

Survey

Thinking about Design

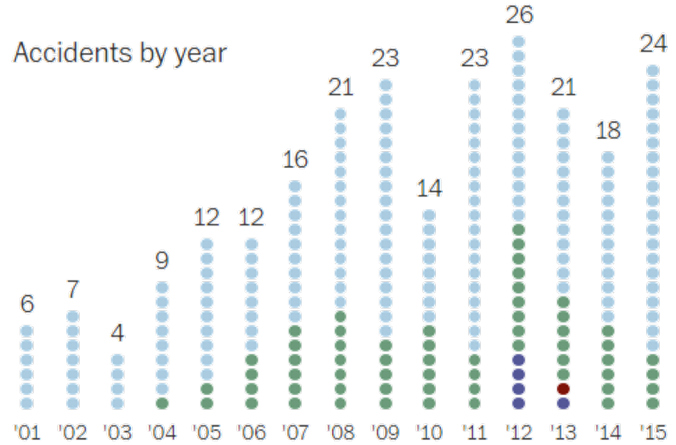
Conclusion

Safety and Privacy: first-class citizens

- Three concerns when talking about build-in:
  - Focus on the systems, from meta-architecture perspective
  - Design strategies, from the design perspective
  - Build-in countermeasures

Autonomous Unmanned Systems

- systematic integration of perception, decision-making, communication, and actuation, embodying machine capabilities to handle the complexity of real world
- without the need of any direct assistance or intervention by human
- Current, efforts are on an ad hoc basis. A systematically taxonomy for safer autonomy



The number of U.S. military drone crashes


Self-driving Uber kills Arizona woman

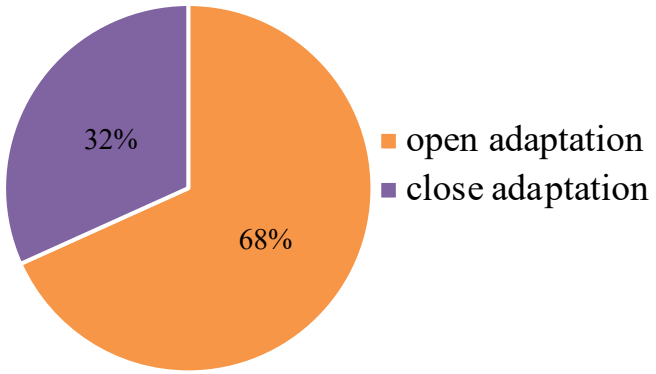
Safety in General

from issue and measure perspective

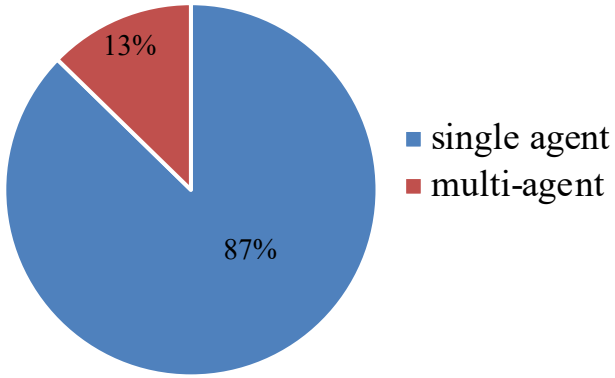
- What does people say:
  - Safety in software and computing system
    - "safety is freedom from accidents or losses" [Leveson 1995]
  - Safety in robots
    - "robot safety involves both issues in the design and implementation of the robot itself, and also issues of human factors relevant to the human-robot interaction" [Graham 1988]
- What do we focus on:
  - Safety threat: undesired or unplanned events bringing about side effects to the system or environment
    - Inside the system: failure, error, conflict among agents
    - From environment: collision with obstacles, other vehicles, cyber attacks
    - To environment: harm to human, property, ecosystem
  - Countermeasure/Solution: to eliminate safety threat or bring it to an acceptable level

Survey Results: Count the Papers

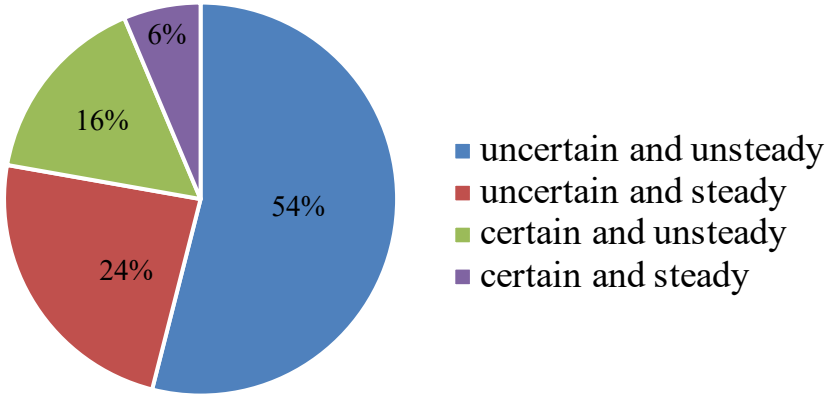
Adaptation mode



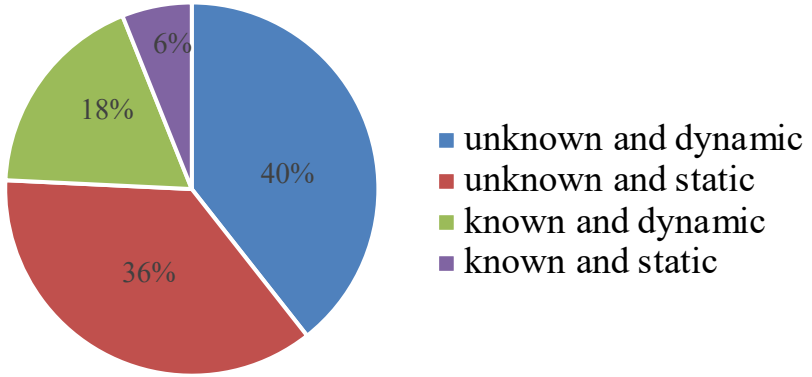
System architecture



Runtime scenarios



Obstacles in Environment

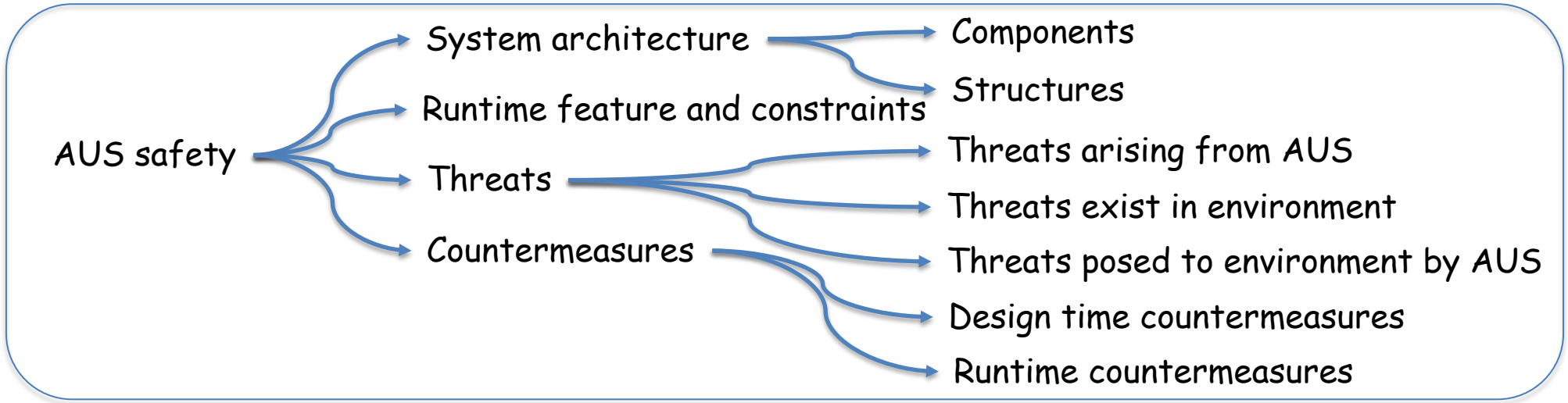


Survey Questions

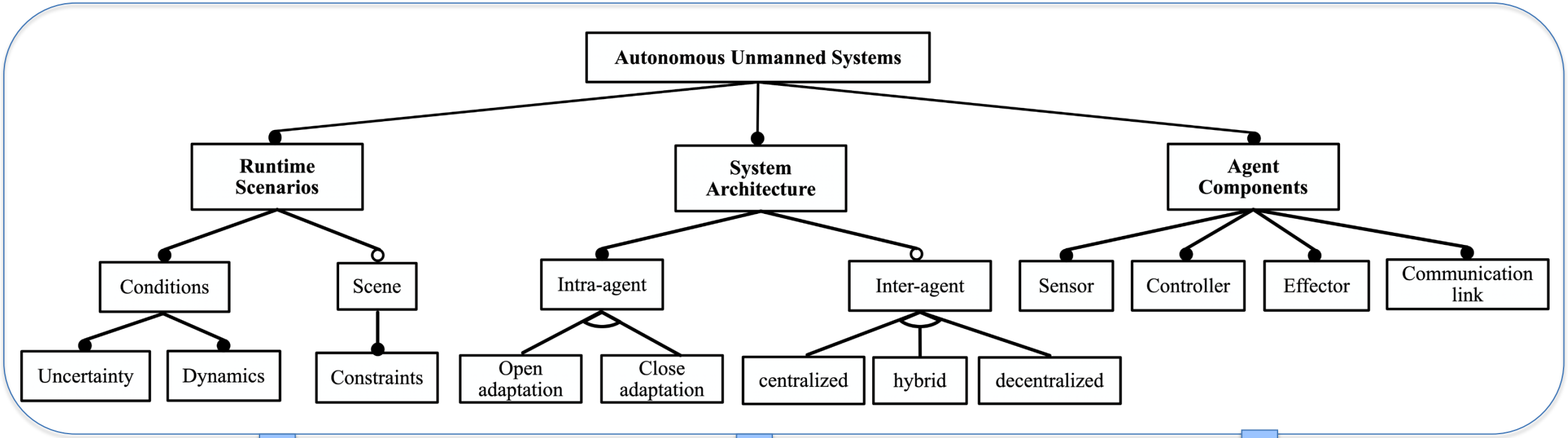
- IEEE Xplore, ACM Digital Library, Science Direct, Springer, and Google Scholar. The search term was "Autonomy" OR "Unmanned" AND ("drone" OR "aircraft" OR "vehicle" OR "vessel" OR "robot" OR "agent" OR "system") AND ("Safety" OR "Risk" OR "Hazard" OR "Critical" OR "Cyber Attack")
- investigate 63 journal and conference papers from 2009 to 2019 about safety concerns for AUSs including drones and other robotic systems.

- RQ1. What are the observations about system architecture?
- RQ2. What are observations at runtime and external constraints?
- RQ3. What are the safety concerns at run-time?
- RQ4. What are the counter-measures?

High Level Taxonomy



Architecture related Concerns

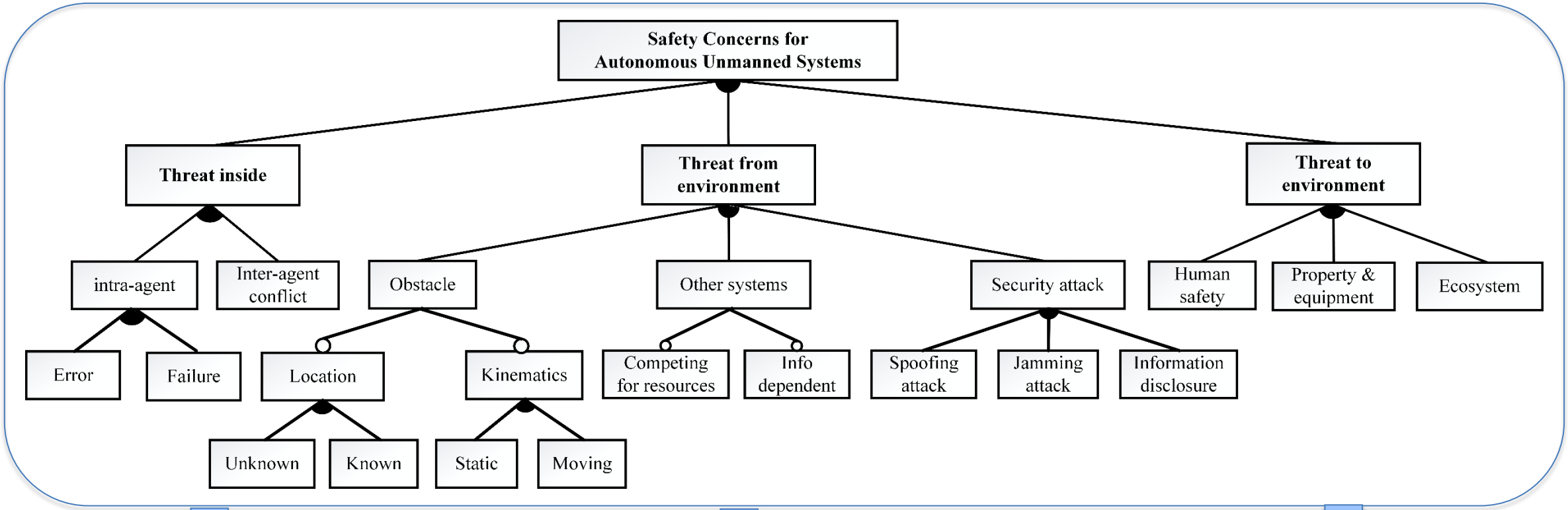


Environment Features

Architecture Features

System Components

Safety Threads at Run-time

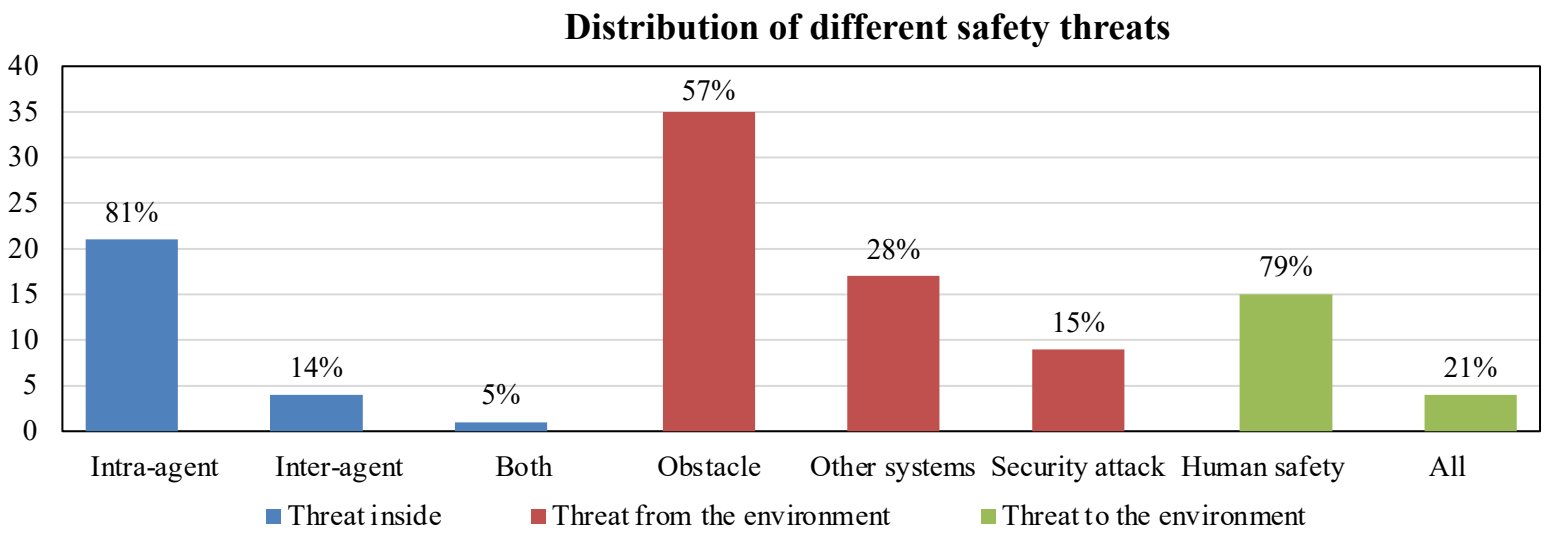


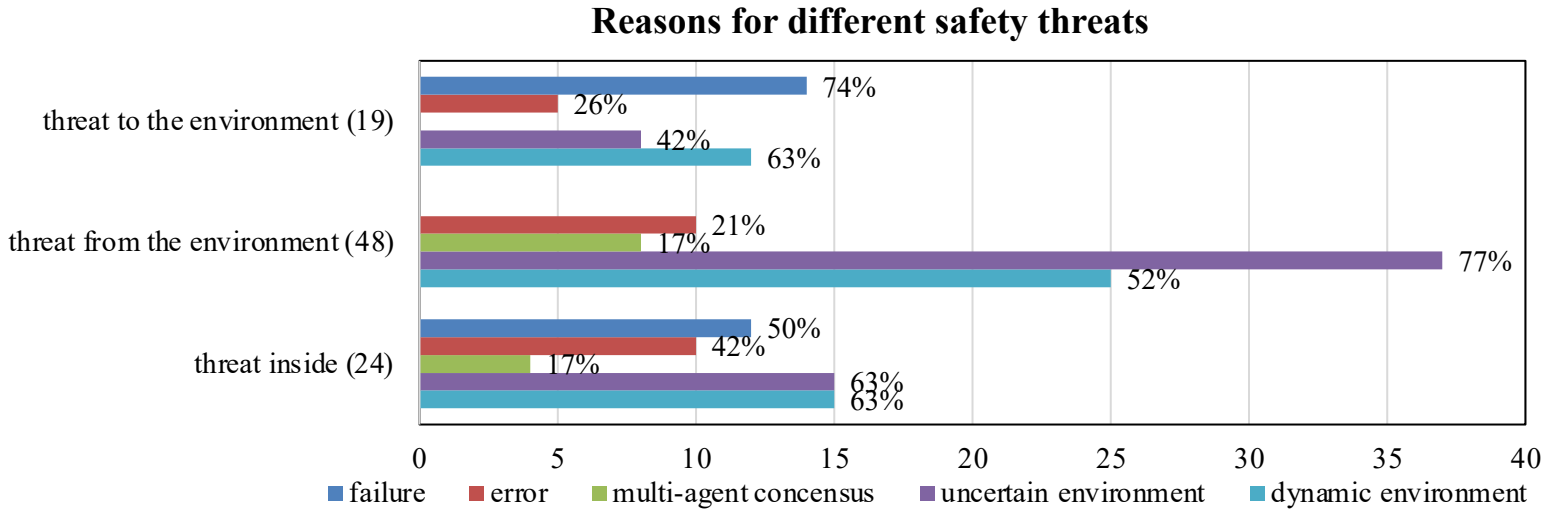
Internal Threads

Threads from External

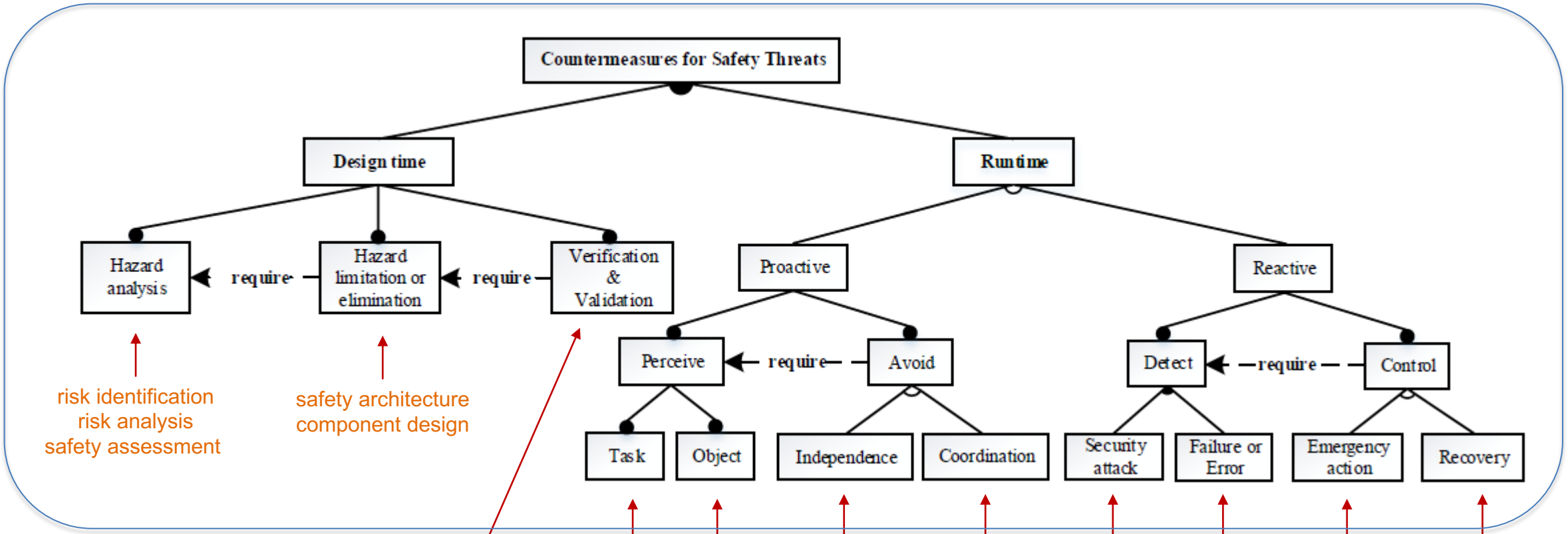
Threads to External

Survey Results: Distributions

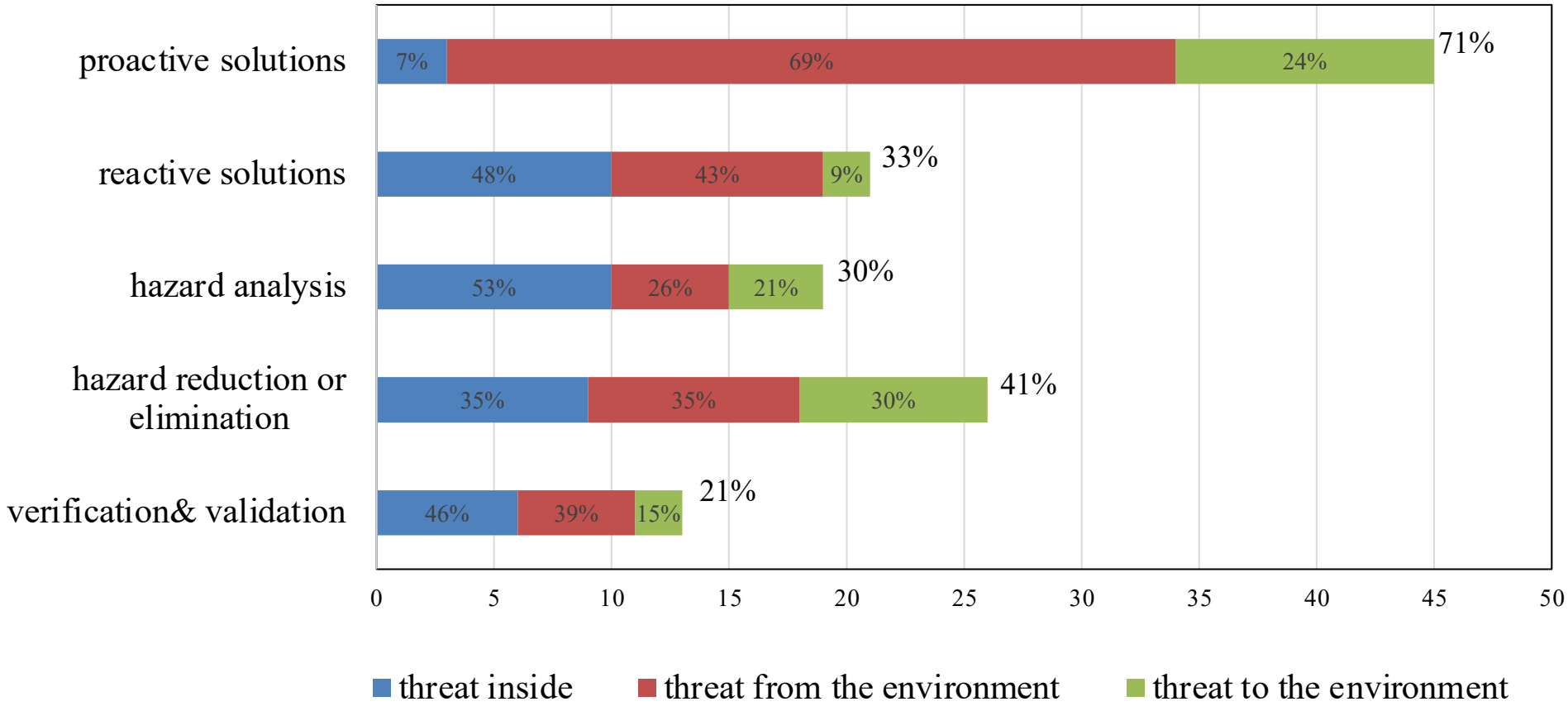




Countermeasures for Safety Threats



Solutions for different safety threats



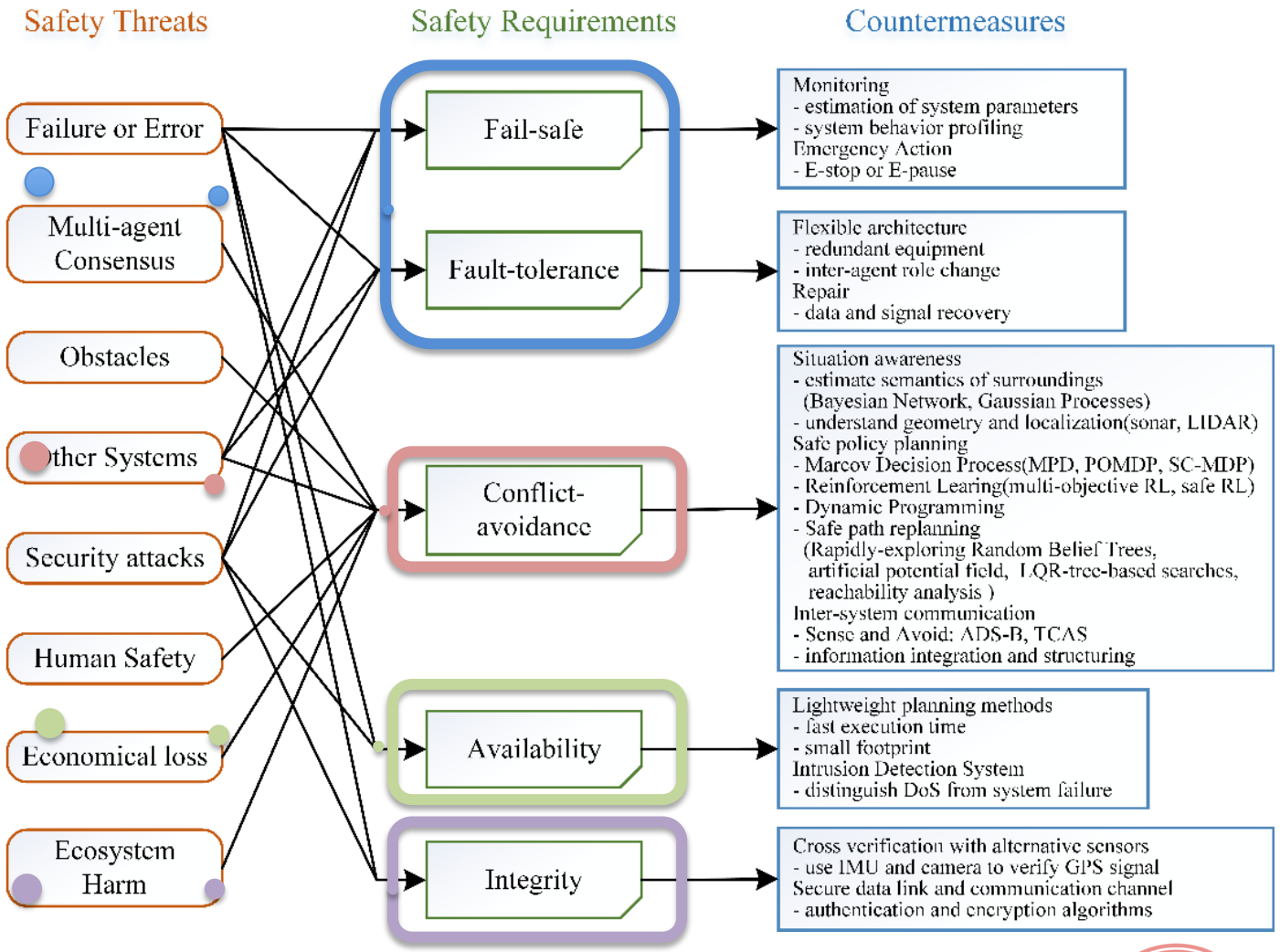
Safety designs for AUS

System Level

Interaction Level

Component Level

Content Level



Environment Features

Table 1. Classifications of application scenarios for autonomous unmanned systems.

Environmental conditions		*Main scene*	*Sub scene*	*System type*	*Representative examples*
Information	Dynamics				
Certain	Static	Land	Indoor	robot	[66]
		Airspace	Low airspace	UAV, drone	[74]
	Dynamic	Land	Indoor	robot	[42] [52] [75] [76] [77]
		Airspace	High airspace	UAS	[78] [65] [79]
			Low airspace	UAV	[68] [69] [80]
		Sea	Under sea	AUV	[73]
Uncertain	Static	Land	Indoor	robot	[48] [81]
			Outside	vehicle	[43] [55] [56]
		Airspace	Low airspace	UAV, drone, quadrotor	[45] [51] [34] [49] [54] [59]
		Sea	Under sea	AUV	[71] [72]
	Dynamic	Land	Indoor	robot	[46] [53] [82]
			Outside	vehicle, robot	[60] [50] [56] [67] [83] [84] [85] [86]
		Airspace	Low airspace	UAV	[70] [62] [87] [58] [88] [40]
			High airspace	UAV, UAS	[89] [90] [91] [64] [33] [92] [93] [94]
		Sea	Above sea	ASV	[95]

Maturity of Existing Solutions

Environment-Centric Safety Requirements		Procedures for safety in AUS															
		Monitoring				Analysing				Planning				Execution			
		proactive		reactive		environment model given		environment model learning		offline		online		independent		collaborative	
Hazard-elimination	fail-safe	○	[30]	●	[26]	●	[27]	○		●	[29]	○	[22]	●	[29]	○	
	fault-tolerance	○	[30]	◑	[27]	●	[30]	○		●	[23]	◑	[32]	●	[33]	◑	[23]
Conflict-avoidance	inter-system	○	[25]	◑	[35]	◑	[34]	○	[44]	◑	[35]	○	[34]	○	[34]	◑	[49]
	human	●	[37]	○	[8]	◑	[8]	◑	[37]	○	[8]	◑	[38]	◑	[38]	○	[39]
	constraints	●	[40]	◑	[43]	●	[24]	◑	[19]	●	[41]	◑	[19]	●	[19]	○	

● Productivity ◑ Enlightenment ○ Trigger

Open Problems and Challenges

Emerging Safety-Critical Scenarios

- Less degrees of fault tolerance
- Privacy of human beings
- Human-drone interaction
- Protected areas
- Forensic required

Environment becomes first citizen

Perception and Detection

- Inter-system collision and conflicts perception
- Perception under uncertain or disturbed environment
- Cyber attacks detection

Advanced Techniques and Algorithms

Platforms for Experimentation

- Simulated environments
- Real-world scenarios

New simulators or test beds

Trade-offs between Safety and Efficiency

- Time efficiency
- Computation efficiency

Design trade-off, finest grain protection

Outline

Motivation

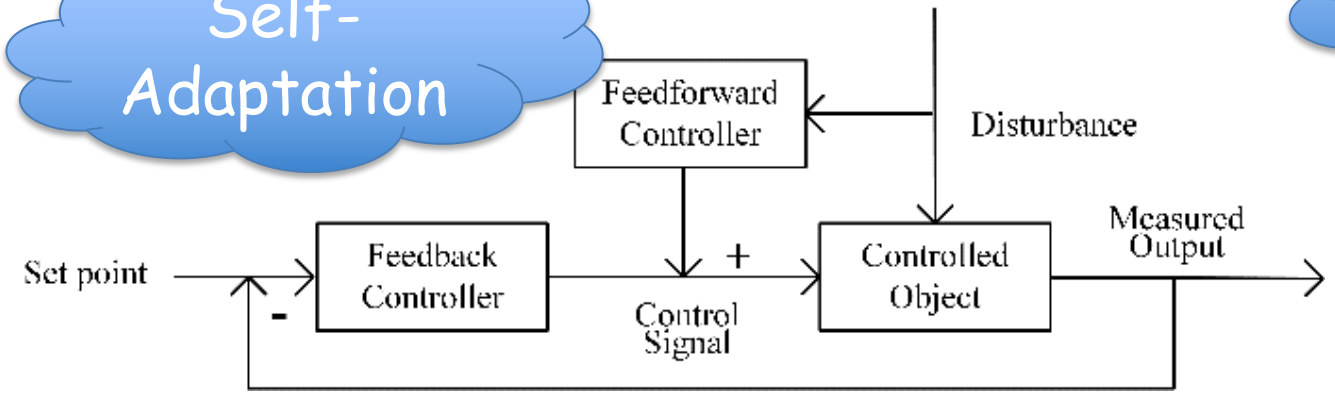
Survey

Thinking about Design

Conclusion

Control Loop and MAPE-K

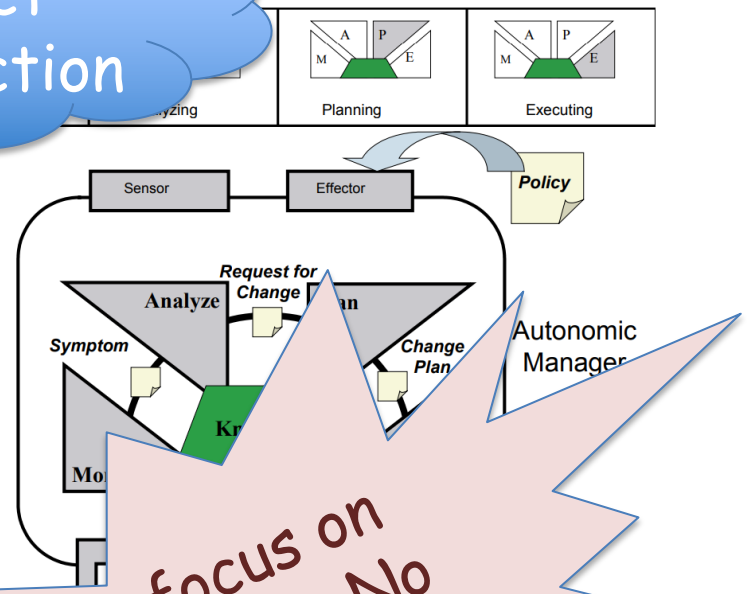
Self-Adaptation



- Feedforward controller: takes into account the external *Disturbances* to produces a *Control Signal* that compensates for the *Disturbances*
- Feedback controller: computes *Control Signal* based on the deviation between desired goals and corresponding *Measured Output*

[Shevtsov 2017]

Direct Interaction



Only focus on the system. No place for environment

Self-ad

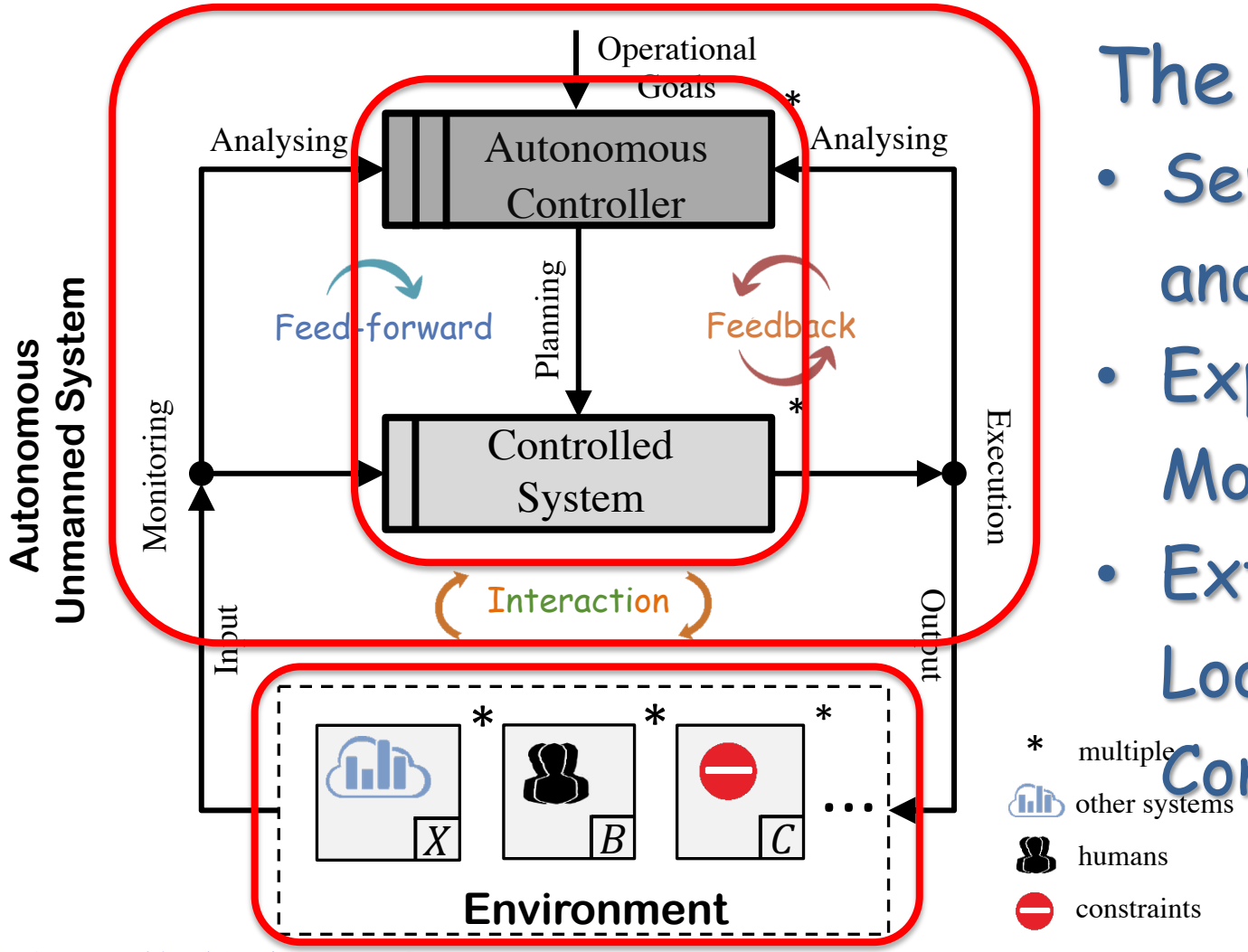
- M

er, effector,

co

2006, Lemos et al. 2010]

Strategy (1) : Architecture: Environment Modeling + Multi-Layer Controls



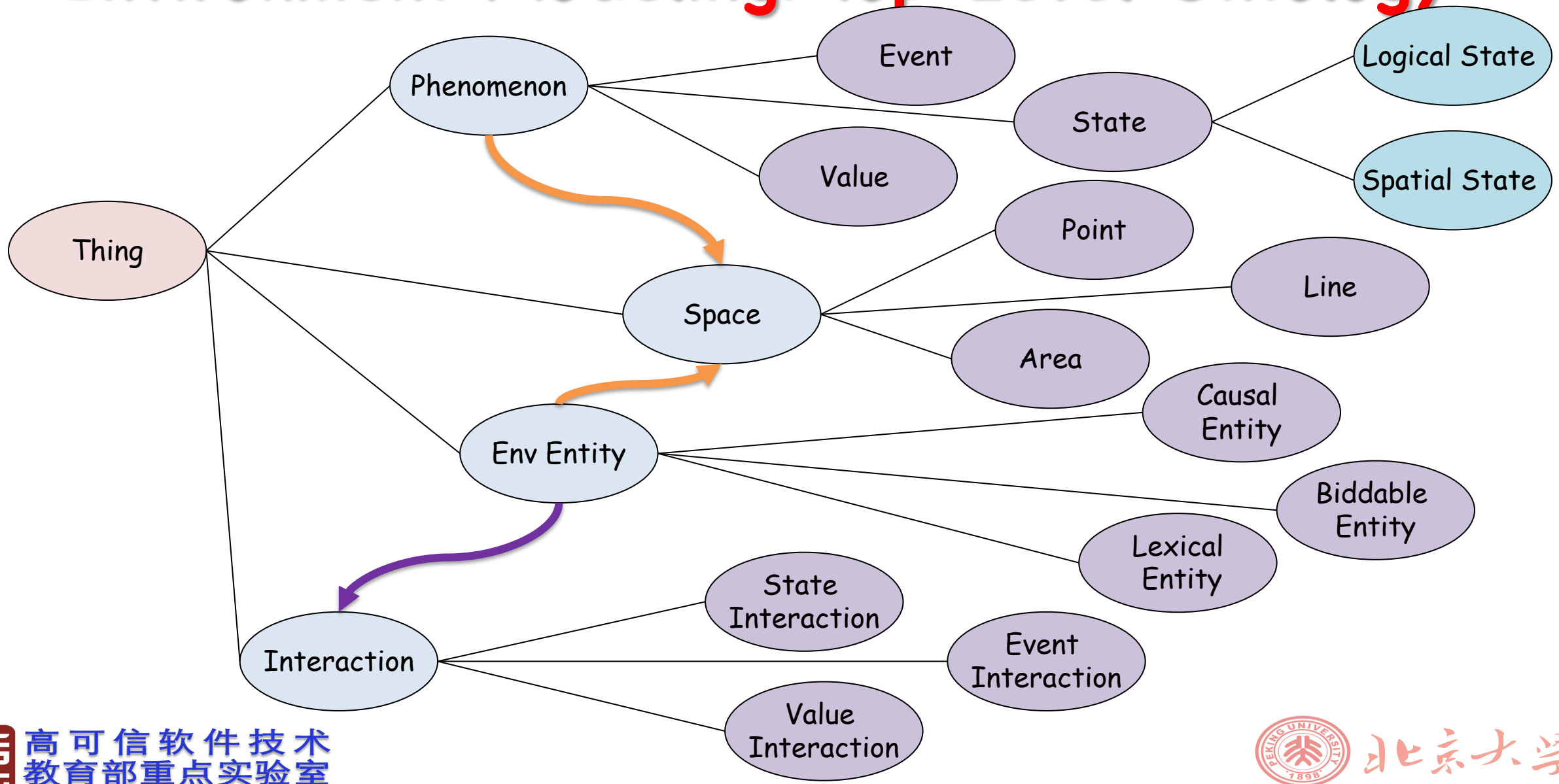
- The fe
- Separation of Core System and Controller
- Explicit
- Modeling
- Extended Unified Control Loops t
- Controls

Unified Control Loops

Separation of Services and Controllers

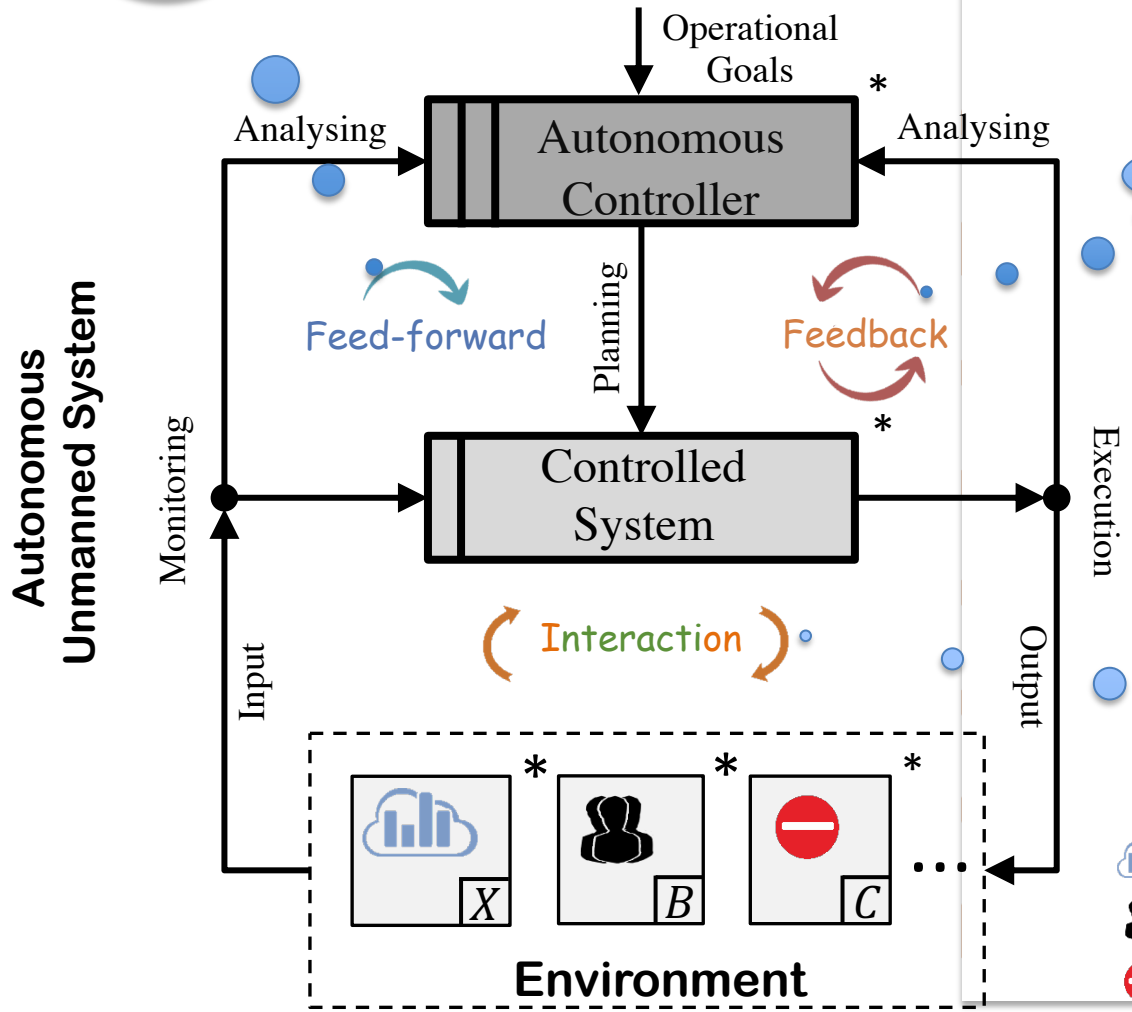
Explicit Environment Modeling

Strategy (2) : Environment Modeling: Top-Level Ontology



Strategy (3) : Fine-Grain Protection

Situational Aware and Adaptation



System Deviation and Risk Control

Environment Dynamics and Uncertainty

* multiple other systems

* humans

* constraints

Countermeasures

Monitoring
- estimation of system parameters
- system behavior profiling
Emergency Action
- E-stop or E-pause

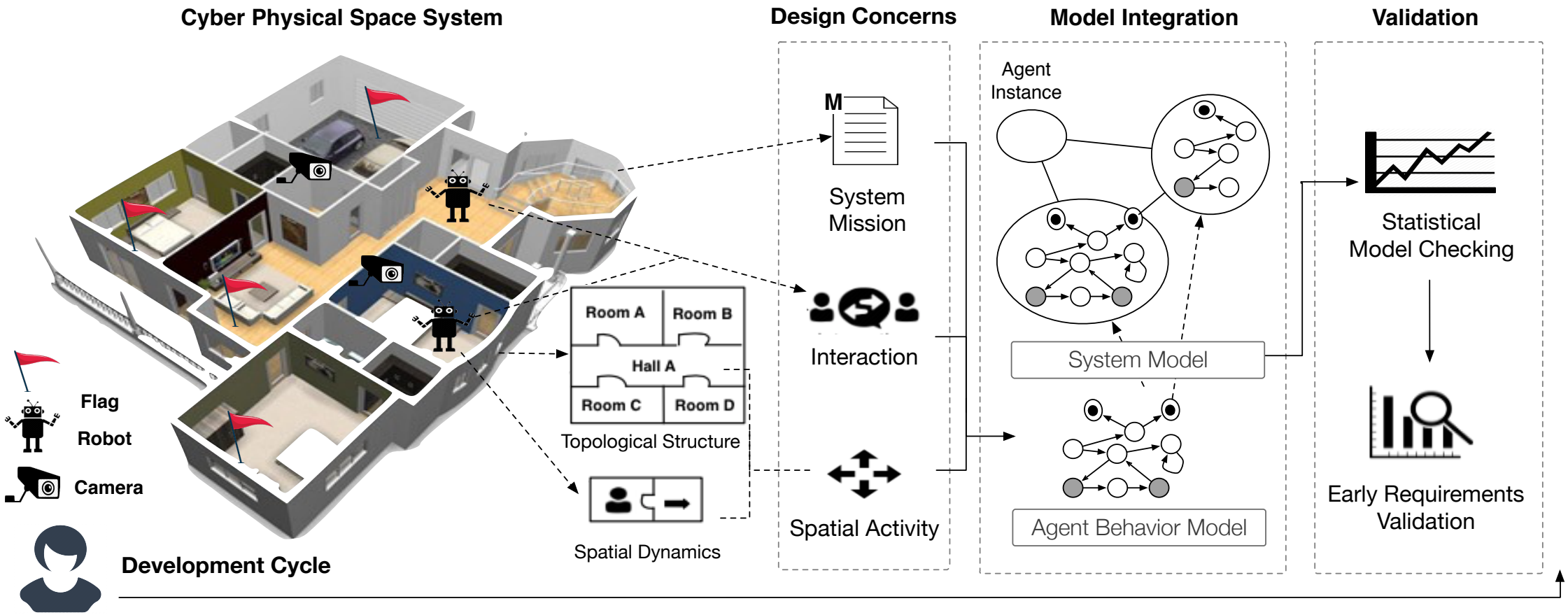
Flexible architecture
- redundant equipment
- inter-agent role change
Repair
- data and signal recovery

Situation awareness
- estimate semantics of surroundings
(Bayesian Network, Gaussian Processes)
- understand geometry and localization(sonar, LIDAR)
Safe policy planning
- Marcov Decision Process(MPD, POMDP, SC-MDP)
- Reinforcement Learing(multi-objective RL, safe RL)
- Dynamic Programming
- Safe path replanning
(Rapidly-exploring Random Belief Trees, artificial potential field, LQR-tree-based searches, reachability analysis)
Inter-system communication
- Sense and Avoid: ADS-B, TCAS
- information integration and structuring

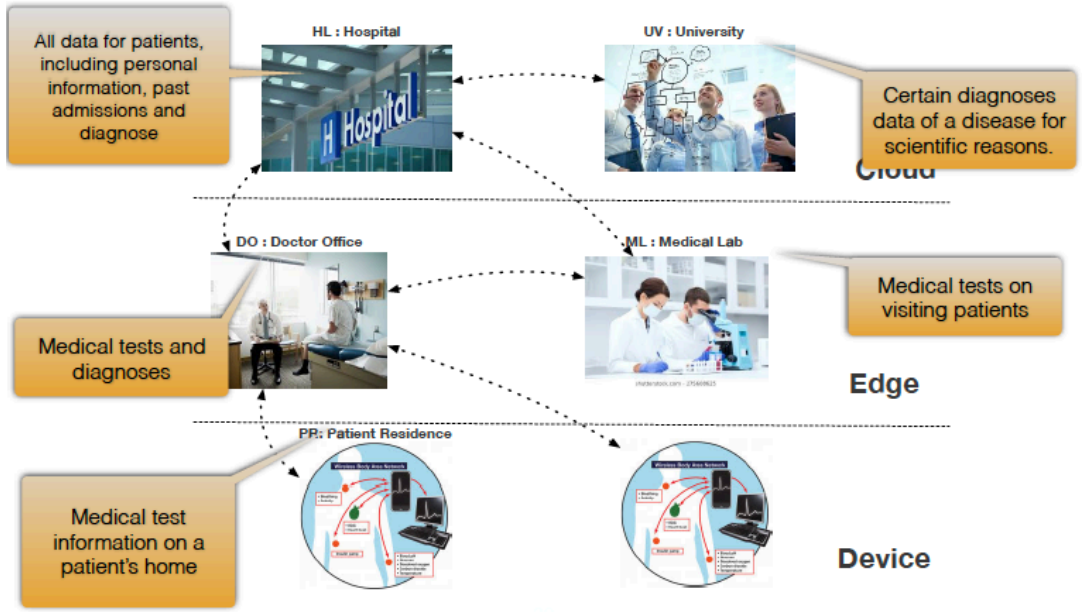
ightweight planning methods
- fast execution time
- small footprint
Intrusion Detection System
- distinguish DoS from system failure

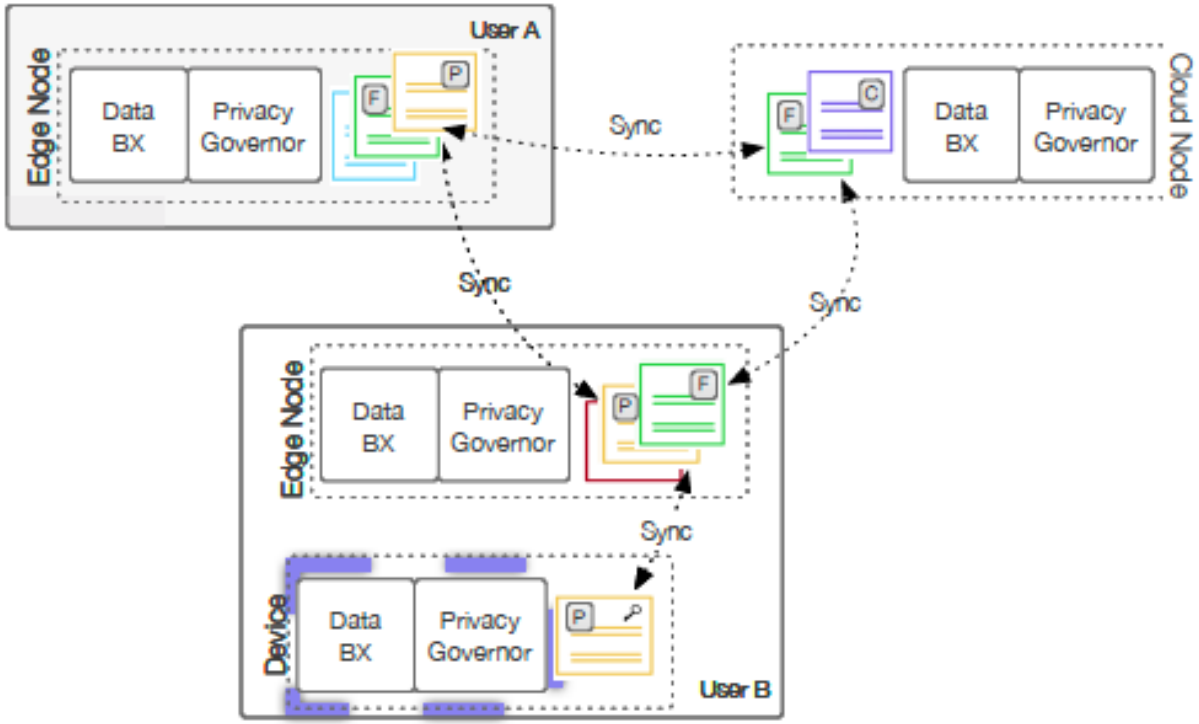
Cross verification with alternative sensors
- use IMU and camera to verify GPS signal
Secure data link and communication channel
- authentication and encryption algorithms

Modeling and Verification on Mission Completion by Moving Agents



Privacy Aware Distributed Data Sharing





Smart Hospital

Runtime Data Sharing Access Control

Risk Aware Motion Planning

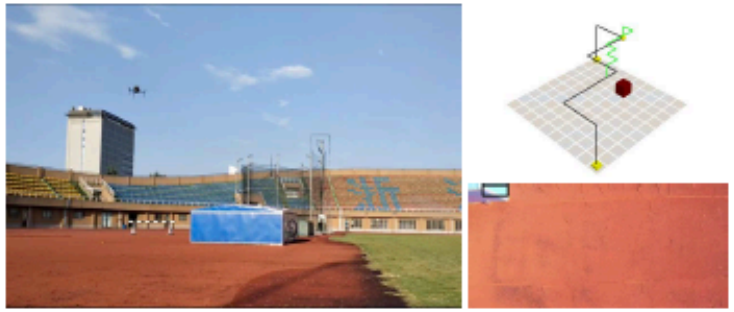
Environment Friendly: Privacy Protection

Runtime Configuration with Scene Detection and Privacy Policy

(a) The first time the private region is detected.

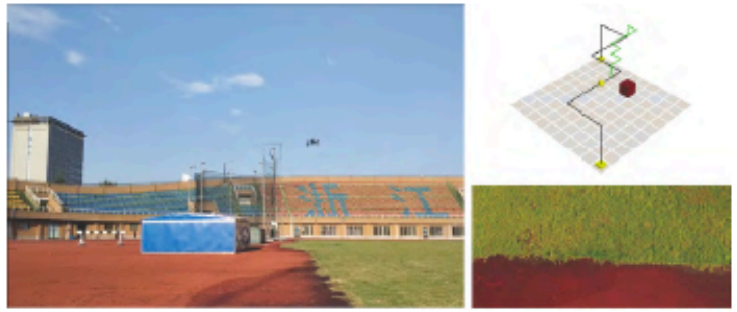
(b) The private region is always in sight.

(c) The UAV passes through the private region.



(a) The UAV detects the private region and re-plans its trajectory and camera orientation.

(b) The private region is avoided by changing trajectory and camera orientation is tuned to 30°.



(c) The UAV passes through the private region, and camera orientation is tuned back to −90°.

Outline

Motivation

Survey

Thinking about Design

Conclusion

More Challenges

- How to identify and model the system bound ?
  - Human intention and domain knowledge
  - Static and dynamic
  - Normal use cases misuse cases, malicious use cases, ……
  - Contract-Based Model
  - ……
- How to deal with the time and space consistency among the environment entities in the real world
  - Space: The space and motion of communicating agents, Robin Milner
  - Time: Real time, time constraints, …….
  - Space and time ?
  - ……

Conclusions and Future Work

- Tighter interaction with the reality brings in the complexity of solution
  - Mixture of continuous and discrete components
  - Discretization and verification of the hybrid models
- Mobility of moving entity along the space topology means the location or context changes
  - Changes of the relationship
  - Location-aware smartness
  - Prediction of the changes
- Environment-friendly, endogenous safety and security
  - Environment risk assessment before taking actions

Messages to Community

- System science and engineering
- Control theory
- Hybrid system modeling and analysis
- Human intention and end-user value traceability
- Constructive components and machine learning components
- ……

- Do the techniques for software analysis, evolution and re-engineering need to be innovated ?

Selected Publications

- Zhi Jin, Environment Modeling-based Requirements Engineering for Software Intensive Systems, Elsevier, Morgan Kaufmann Publisher, 2018
- Nianyu Li, Christos Tsigkanos, Zhi Jin, Zhenjiang Hu and Carlo Ghezzi, Early validation of cyber-physical space systems via multi-concerns integration, Journal of System and Software 170: 110742, 2020
- Lionel Montrieux, Naoyasu Ubayashi, Tianqi Zhao, Zhi Jin and Zhenjiang Hu, Bidirectional Transformations for Self-Adaptive Systems, Communications of NII Shonan Meetings, Engineering Adaptive Software System 2019: 95-114, Springer, 2019
- Nianyu Li, Christos Tsigkanos, Zhi Jin, Schahram Dustdar, Zhenjiang Hu and Carlo Ghezzi, POET: Privacy on the Edge with Bidirectional Data Transformations, 2019 IEEE International Conference on Pervasive Computing and Communications (PerCom 2019):
- Tianqi Zhao, Wei Zhang, Haiyan Zhao, Zhi Jin: A Reinforcement Learning-Based Framework for the Generation and Evolution of Adaptation Rules. ICAC 2017: 103-112
- Yixing Luo, Yijun Yu, Zhi Jin, Haiyan Zhao, Environment-Centric Safety Requirements for Autonomous Unmanned Systems, RE@NEXT 2019
- Christos Tsigkanos, Nianyu Li, Zhi Jin, Zhenjiang Hu, Carlo Ghezzi, Scalable Multiple-View Analysis of Reactive Systems via Bidirectional Model Transformations, Proceedings of the 35th IEEE/ACM International Conference on Automated Software Engineering, 2020 (ACM Distinguished Paper Award)
- Yixing Luo, Yijun Yu, Zhi Jin, Yao Li, Zuohua Ding, Yuan Zhou, Yang Liu: Privacy-Aware UAV Flights through Self-Configuring Motion Planning. ICRA 2020: 1169-1175

Acknowledgements

- Key Projects of National Natural Science Foundation of China under Grant Nos. 90818026, 61620106007, 61751210
- National Grand Fundamental Research Program of China under Grant No. 2009CB320701, Ministry of Science and Technology

国家自然科学基金委员会 National Natural Science Foundation of China

中华人民共和国科学技术部 Ministry of Science and Technology of the People's Republic of China

- Thanks are due to
  - Collaborators: Schahram Dustdar, Carlo Ghezzi, Zhenjiang Hu, Lionel Montrieux, Christos Tsigknos, Naoyasu Ubayashi, Yijun Yu, Haiyan Zhao, Wei Zhang
  - Students: Nianyu Li, Yixing Luo

Thanks
For Your Attentions