Juggling the Jigsaw

Towards Automated Problem Inference from Network Trouble Tickets

Rahul Potharaju (Purdue University)

Navendu Jain (Microsoft Research)

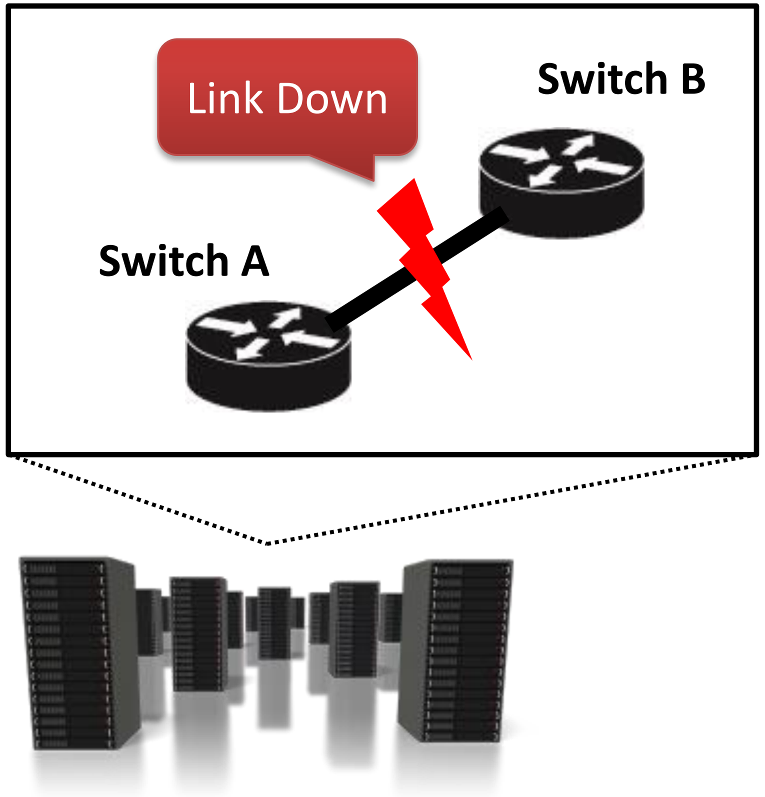
Cristina Nita-Rotaru (Purdue University)

PURDUE UNIVERSITY

April 3, 2013
NSDI 2013

Microsoft® Research

Network Troubleshooting: The Big Picture



Datacenters

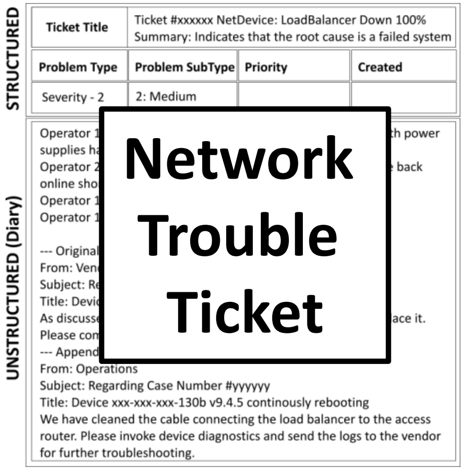
Alert

Network Monitoring

Trigger

Operator Console

Log Ticket



STRUCTURED

Ticket Title	Ticket #xxxxxx NetDevice: LoadBalancer Down 100% Summary: Indicates that the root cause is a failed system		
Problem Type	Problem SubType	Priority	Created
Severity - 2	2: Medium		

UNSTRUCTURED (Diary)

Operator 1 … th power
supplies ha…
Operator 2 … back
online sho…
Operator 1…
Operator 1…

--- Original…
From: Ven…
Subject: Re…
Title: Devic…
As discusse… ace it.
Please con…
--- Append…
From: Operations
Subject: Regarding Case Number #yyyyyy
Title: Device xxx-xxx-xxx-130b v9.4.5 continously rebooting
We have cleaned the cable connecting the load balancer to the access router. Please invoke device diagnostics and send the logs to the vendor for further troubleshooting.

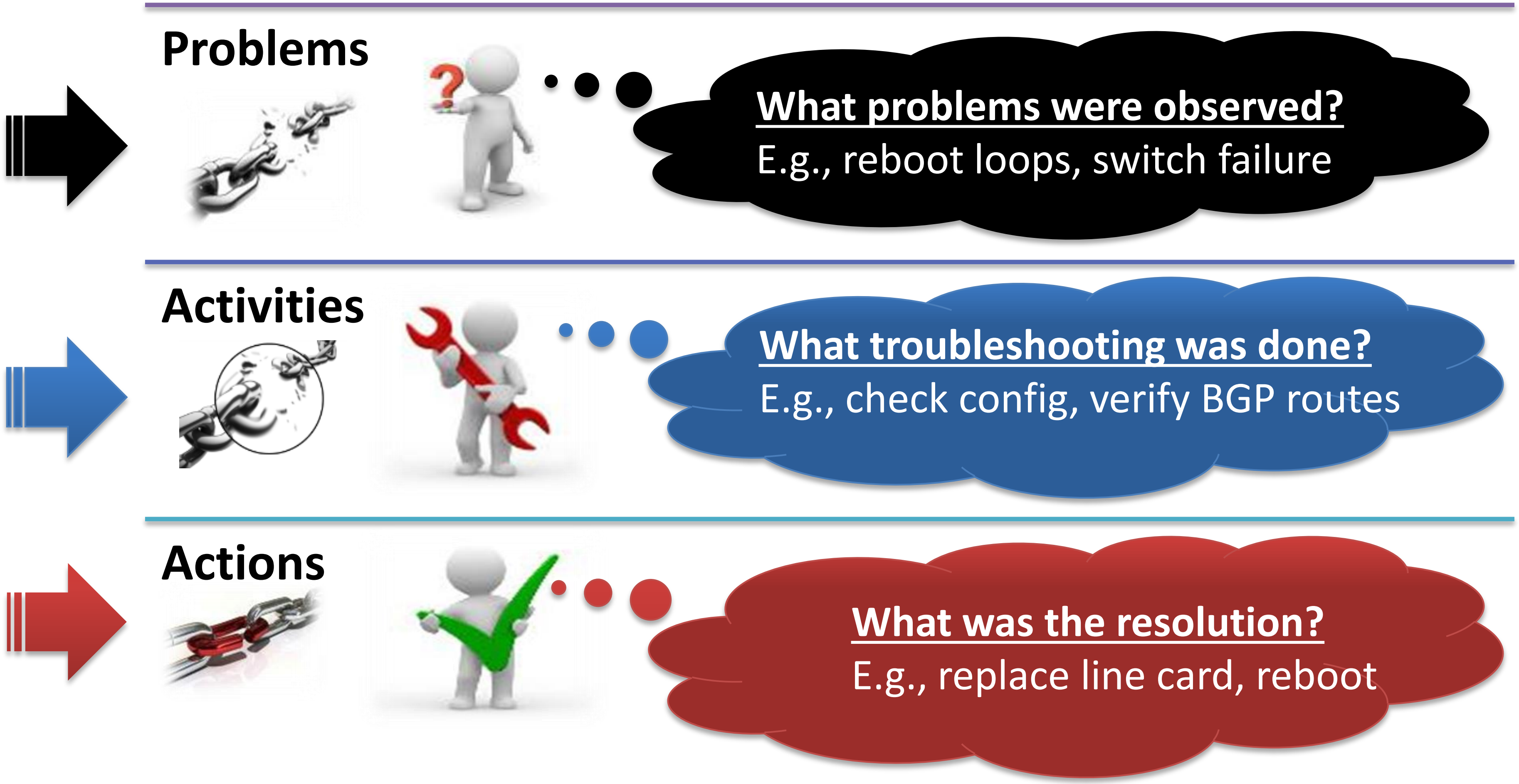
Network Trouble Ticket

Diaries written by operators during network troubleshooting

Goal: Automated Problem Inference from Trouble Tickets

Network trouble ticket

Inference Output



Problems

What problems were observed?
E.g., reboot loops, switch failure

Activities

What troubleshooting was done?
E.g., check config, verify BGP routes

Actions

What was the resolution?
E.g., replace line card, reboot

Goal: Automated Problem Inference from Trouble Tickets

Inference Output

Problems

Key questions for network management

[Q1]: Why is network redundancy ineffective?

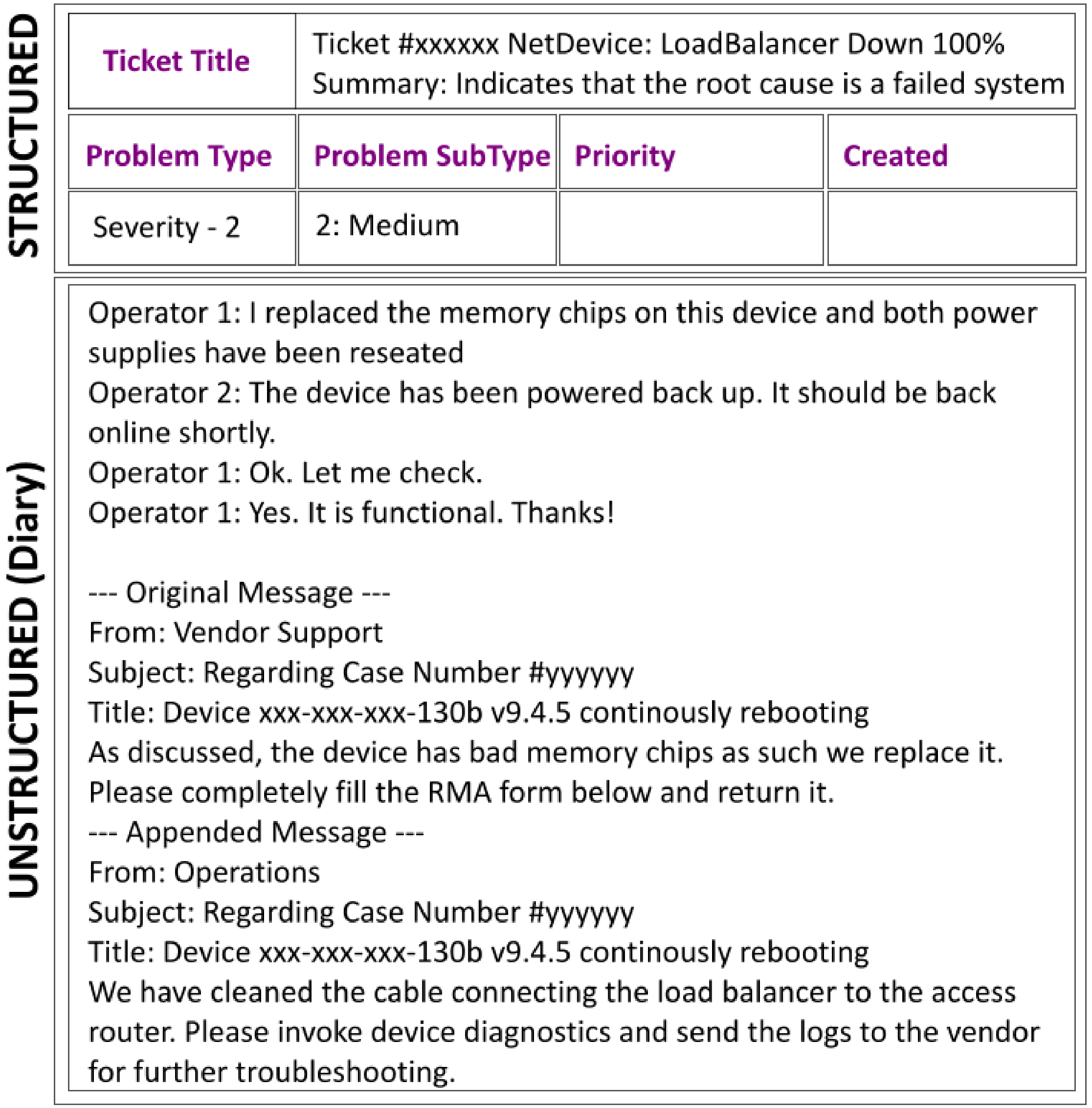
[Q2]: What are the top-k failing components?

[Q3]: Are new devices more reliable?

What was the resolution?

E.g., replace line card, reboot

What Does a Ticket Contain?



STRUCTURED

Ticket Title	Ticket #xxxxxx NetDevice: LoadBalancer Down 100% Summary: Indicates that the root cause is a failed system			
Problem Type	Problem SubType	Priority	Created	
Severity - 2	2: Medium			

UNSTRUCTURED (Diary)

Operator 1: I replaced the memory chips on this device and both power supplies have been reseated
Operator 2: The device has been powered back up. It should be back online shortly.
Operator 1: Ok. Let me check.
Operator 1: Yes. It is functional. Thanks!

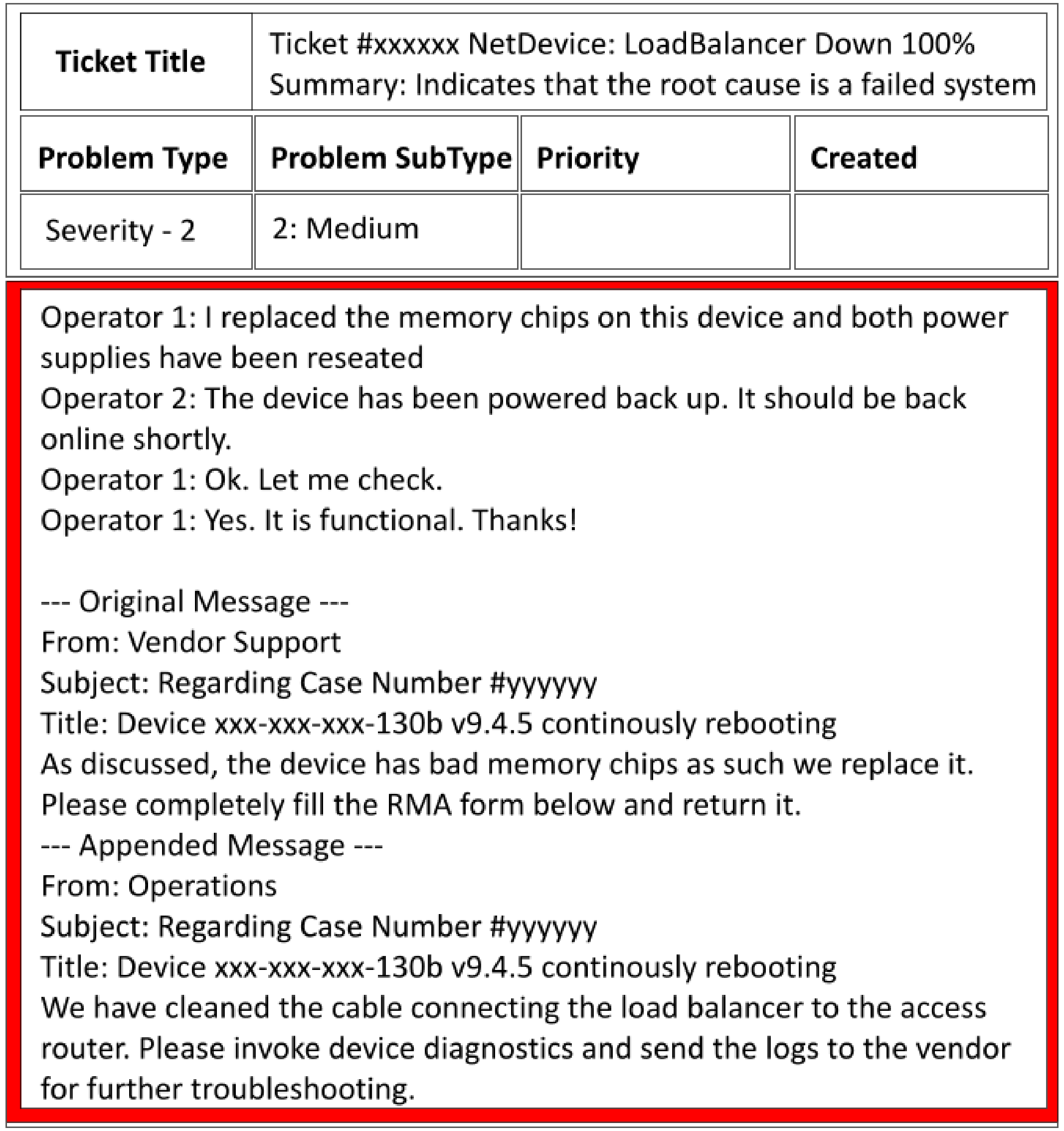
--- Original Message ---
From: Vendor Support
Subject: Regarding Case Number #yyyyyy
Title: Device xxx-xxx-xxx-130b v9.4.5 continously rebooting
As discussed, the device has bad memory chips as such we replace it. Please completely fill the RMA form below and return it.
--- Appended Message ---
From: Operations
Subject: Regarding Case Number #yyyyyy
Title: Device xxx-xxx-xxx-130b v9.4.5 continously rebooting
We have cleaned the cable connecting the load balancer to the access router. Please invoke device diagnostics and send the logs to the vendor for further troubleshooting.

STRUCTURED FIELDS
E.g., ticket title, problem type, priority etc.

FREE-FORM TEXT
E.g., operator notes, emails, device debug logs, etc.

Challenges in Analyzing Trouble Tickets

STRUCTURED



Ticket Title	Ticket #xxxxxx NetDevice: LoadBalancer Down 100% Summary: Indicates that the root cause is a failed system		
Problem Type	**Problem SubType**	**Priority**	**Created**
Severity - 2	2: Medium		

UNSTRUCTURED (Diary)

Operator 1: I replaced the memory chips on this device and both power supplies have been reseated
Operator 2: The device has been powered back up. It should be back online shortly.
Operator 1: Ok. Let me check.
Operator 1: Yes. It is functional. Thanks!

--- Original Message ---
From: Vendor Support
Subject: Regarding Case Number #yyyyyy
Title: Device xxx-xxx-xxx-130b v9.4.5 continously rebooting
As discussed, the device has bad memory chips as such we replace it.
Please completely fill the RMA form below and return it.
--- Appended Message ---
From: Operations
Subject: Regarding Case Number #yyyyyy
Title: Device xxx-xxx-xxx-130b v9.4.5 continously rebooting
We have cleaned the cable connecting the load balancer to the access router. Please invoke device diagnostics and send the logs to the vendor for further troubleshooting.

- Coarse-grained information
- Inaccurate or Incomplete: 69%-75% in 10K+ tickets in our study!

- Written in natural language
- Typos and ambiguity
- Grammatical errors
- Domain-specific terms
  - E.g., DNS, DMZ, line card

Our Contributions

- **Measurement study**: 10K+ tickets logged from a large cloud provider (April 2010-12)
  - Coarse-grained and inaccurate structured data in 69%-75% of the tickets
  - Free-form natural language text comprising emails, IMs, device debug logs, etc.

- **NetSieve**: Combines NLP, knowledge discovery and ontology modeling in a novel way
  1. **Problems**: Network entity and its state/condition e.g., firewall failure, firmware error
  2. **Activities**: Steps performed during troubleshooting e.g., change config, verify routes
  3. **Actions**: Resolution applied to mitigate the problem e.g., replace disk, reboot switch

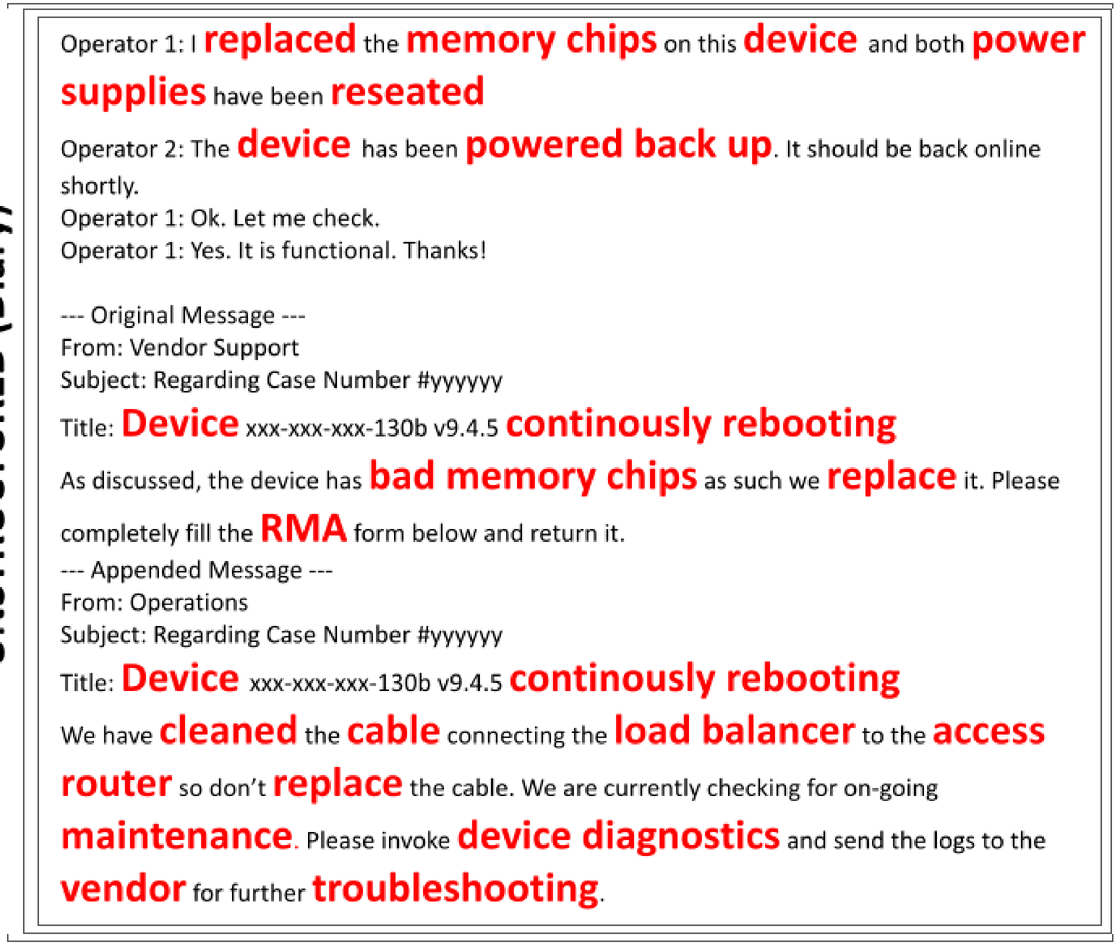
- **Achieves 83%-100% accuracy**
  - Evaluated using a domain-expert, hardware vendor tickets and a survey of operators

Roadmap

- Motivation
- Strawman Approaches to Analyze Free-form Text
- NetSieve: Semantic-based Approach
- Evaluation
- Conclusion

Strawman Approach To Analyze Free-form Text

UNSTRUCTURED (Diary)



Operator 1: I **replaced** the **memory chips** on this **device** and both **power supplies** have been **reseated**

Operator 2: The **device** has been **powered back up**. It should be back online shortly.
Operator 1: Ok. Let me check.
Operator 1: Yes. It is functional. Thanks!

--- Original Message ---
From: Vendor Support
Subject: Regarding Case Number #yyyyyy
Title: **Device** xxx-xxx-xxx-130b v9.4.5 **continously rebooting**
As discussed, the device has **bad memory chips** as such we **replace** it. Please completely fill the **RMA** form below and return it.
--- Appended Message ---
From: Operations
Subject: Regarding Case Number #yyyyyy
Title: **Device** xxx-xxx-xxx-130b v9.4.5 **continously rebooting**
We have **cleaned** the **cable** connecting the **load balancer** to the **access router** so don't **replace** the cable. We are currently checking for on-going **maintenance**. Please invoke **device diagnostics** and send the logs to the **vendor** for further **troubleshooting**.

Strawman #1: Use NLP techniques

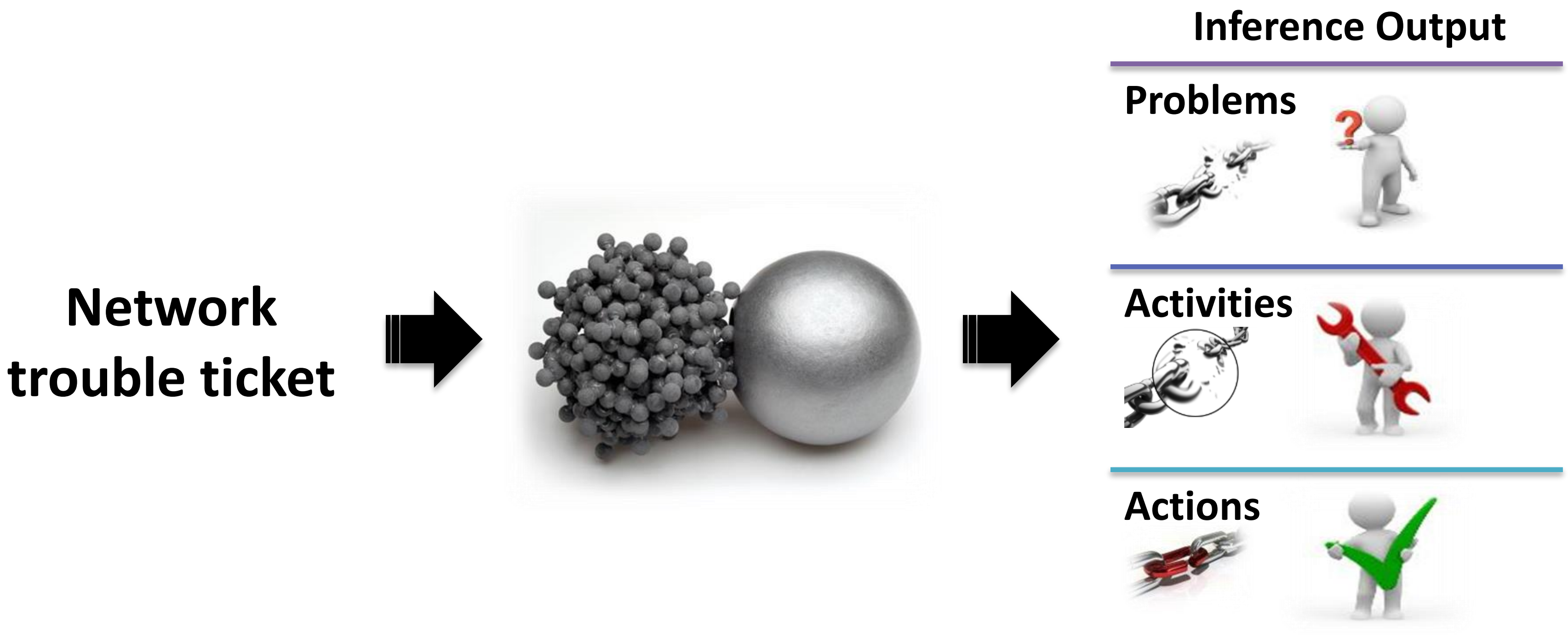
Limitation: Work only on well-written text such as news-articles

Strawman #2: Keyword selection

Limitations: Ignores contextual semantics

Strawman #3: Clustering tickets based on manual keyword selection

Limitations: 1. Significant time and effort to build the keyword list
2. Limited coverage or risks becoming outdated as the network evolves



Network trouble ticket

Inference Output

Problems

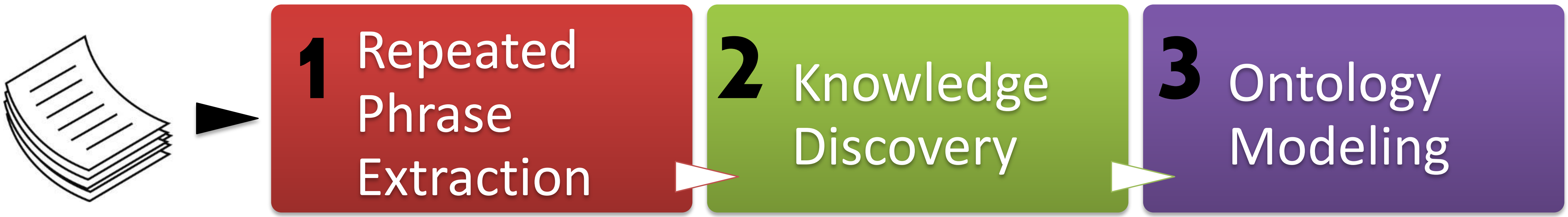
Activities

Actions

NetSieve: Semantic-based Approach to Do Problem Inference

NetSieve Architecture

KNOWLEDGE BUILDING PHASE



TROUBLE TICKET REPOSITORY

Goal: Find frequently occurring phrases

Goal: Find phrases important in the “networking” domain

Goal: Semantic interpretation of the domain-specific phrases

… power supply unit is faulty…
… access router inoperative…
… run config script …
… is to inform you that there …

<power supply unit is faulty>
<access router inoperative>
<run config script>

- ENTITY: power supply unit -> STATE: faulty
- ENTITY: access router -> CONDITION: inoperative
- ENTITY: config script -> ACTION: run

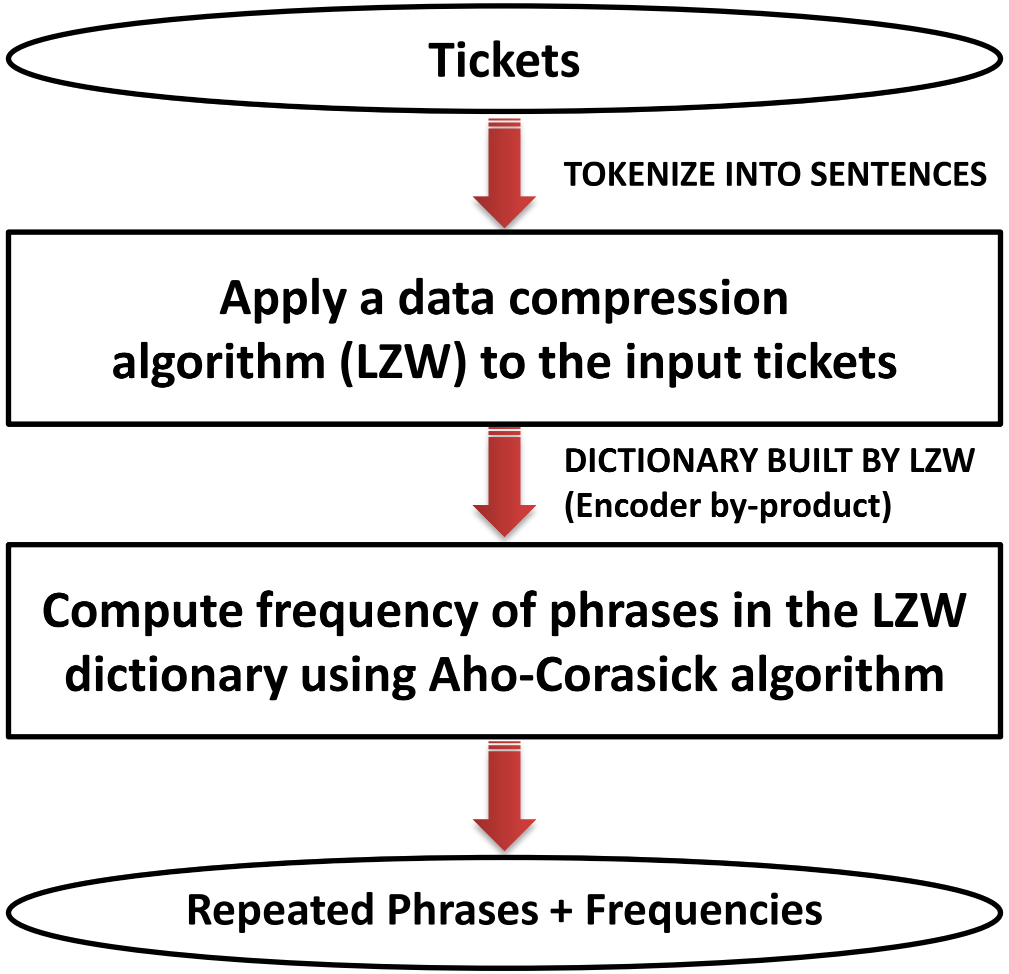
1 Repeated Phrase Extraction

2 Knowledge Discovery

3 Ontology Modeling

Step – I: Repeated Phrase Extraction

- **Goal**: Find frequently occurring phrases
  - Extracting all possible n-grams
- **Challenges**:
  - Computationally expensive
  - Fine-tuning numerous thresholds
  - Not all n-grams are useful (noise)
- **Approach**: Trade completeness for speed and scalability


Tickets

TOKENIZE INTO SENTENCES

Apply a data compression algorithm (LZW) to the input tickets

DICTIONARY BUILT BY LZW (Encoder by-product)

Compute frequency of phrases in the LZW dictionary using Aho-Corasick algorithm

Repeated Phrases + Frequencies

1 Repeated Phrase Extraction

2 Knowledge Discovery

3 Ontology Modeling

Step – II: Knowledge Discovery

- **Goal**: Find phrases <u>important</u> in the current domain to do problem inference

- **Challenges**:
  - Filter meaningful phrases from noisy ones
  - Expert-labeling is time-consuming

- **Approach: (19M phrases → 5.6K phrases)**
  1. Apply a pipeline of linguistic filters
  2. Rank phrases by importance using information theoretic measures

Phrase	Important?
power disruption on access router	✓
key corruption due to expired certificate	✓
bad memory on server	✓
prior communication	✗
best regards	✗
informing you that	✗

Step – III: Ontology Modeling

- **Goal**: Semantic interpretation of the extracted important phrases in the current domain

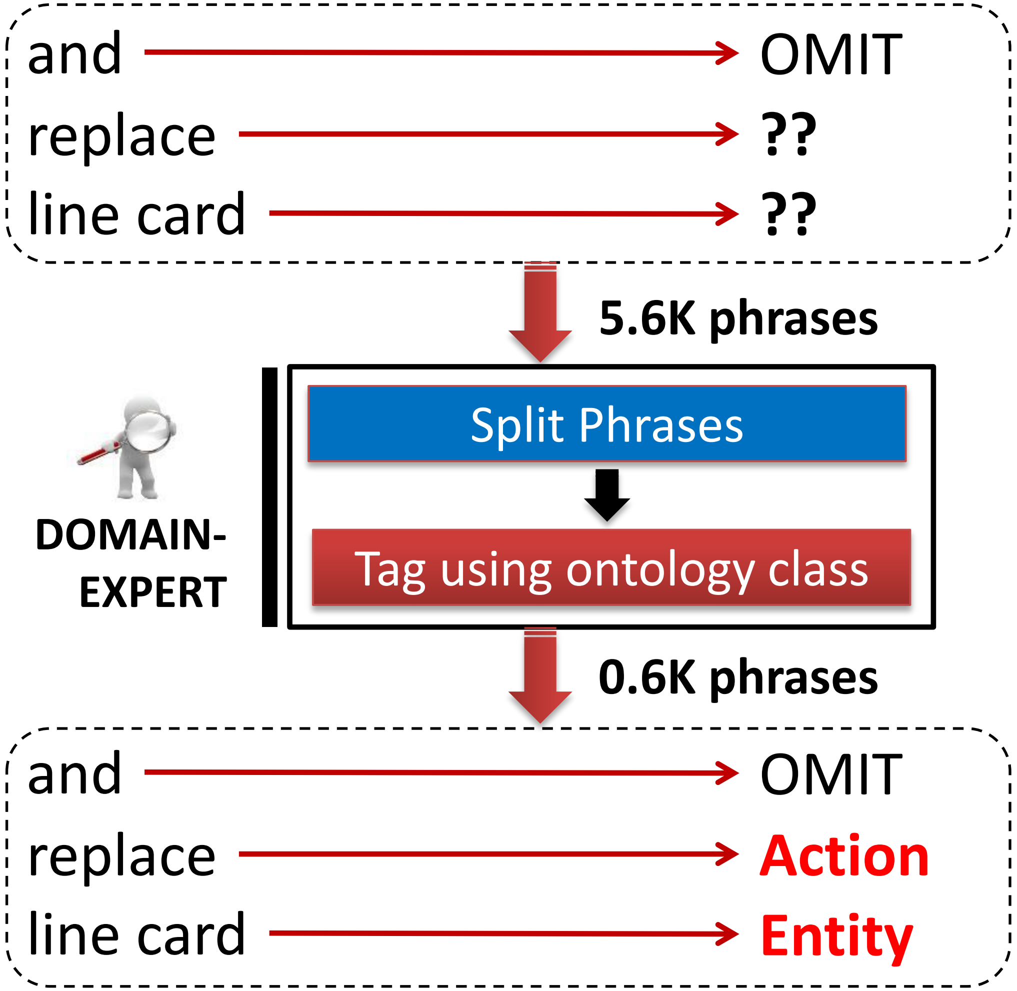
- **Challenges**:
  – How to precisely define the meaning of domain-specific phrases and relationships between them?

- **Approach**:
  1. Define an ontology model
  2. Tag phrases with classes from the ontology model

and → OMIT
replace → ??
line card → ??

Step – III: Ontology Modeling

	Semantic Meaning
Entity	Object that can be deployed or repaired e.g., flash memory, core router
Action	Behavior that can be caused upon an entity e.g., reboot, replace
Condition	Describes the state of an entity e.g., bit error, hung state
…	…



and → OMIT
replace → ??
line card → ??

5.6K phrases

DOMAIN-EXPERT

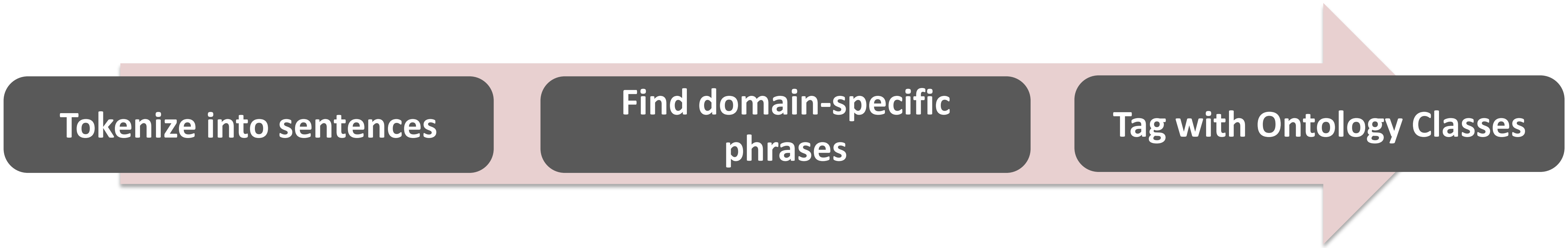
Split Phrases

Tag using ontology class

0.6K phrases

and → OMIT
replace → **Action**
line card → **Entity**

Putting it All Together (1/2): Tagging a Ticket



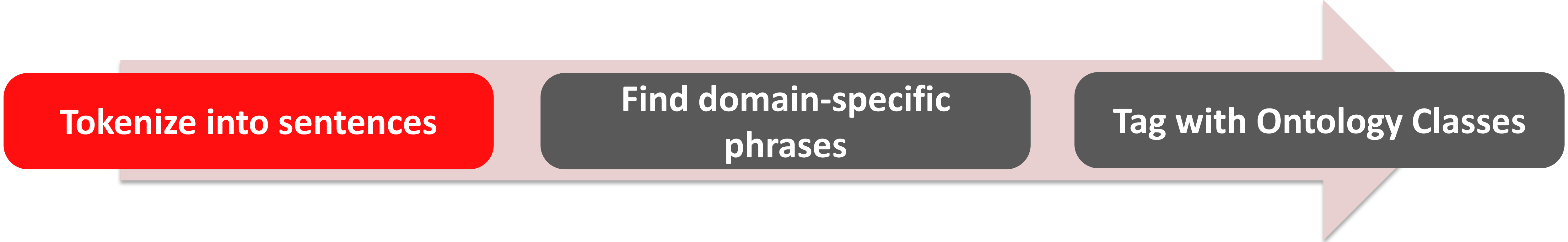
Tokenize into sentences

Find domain-specific phrases

Tag with Ontology Classes

We have raised a request #9646604 and found that the device xxx-xxx-xxx-130a Power LED is amber and it is in hung state. We checked the device for connectivity issues, cleaned the fiber and found that the power supply unit is faulty. We replaced the power supply unit.

Putting it All Together (1/2): Tagging a Ticket

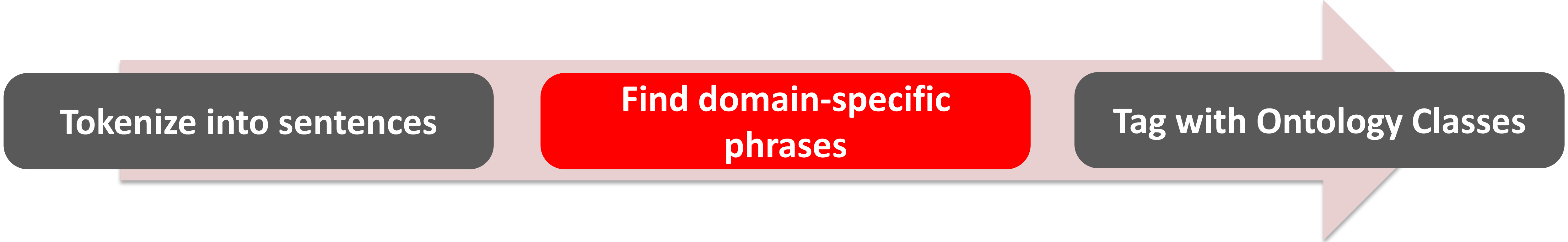


We have raised a request #9646604 and found that the device xxx-xxx-xxx-130a Power LED is amber and it is in hung state.

We checked the device for connectivity issues, cleaned the fiber and found that the power supply unit is faulty.

We replaced the power supply unit.

Putting it All Together (1/2): Tagging a Ticket



Tokenize into sentences

Find domain-specific phrases

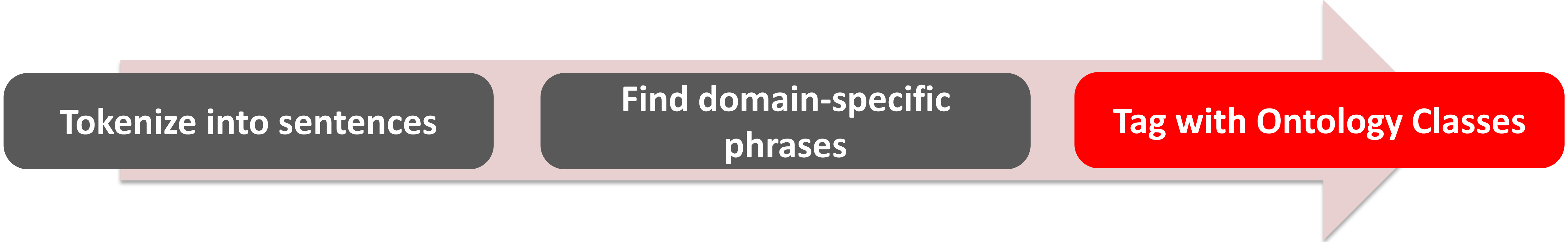
Tag with Ontology Classes

We have raised a request #9646604 and found that the **device** xxx-xxx-xxx-130a **Power LED** is **amber** and it is in **hung state**.

We **checked** the **device** for **connectivity issues**, **cleaned** the **fiber** and found that the **power supply unit** is **faulty**.

We **replaced** the **power supply unit**.

Putting it All Together (1/2): Tagging a Ticket



Tokenize into sentences

Find domain-specific phrases

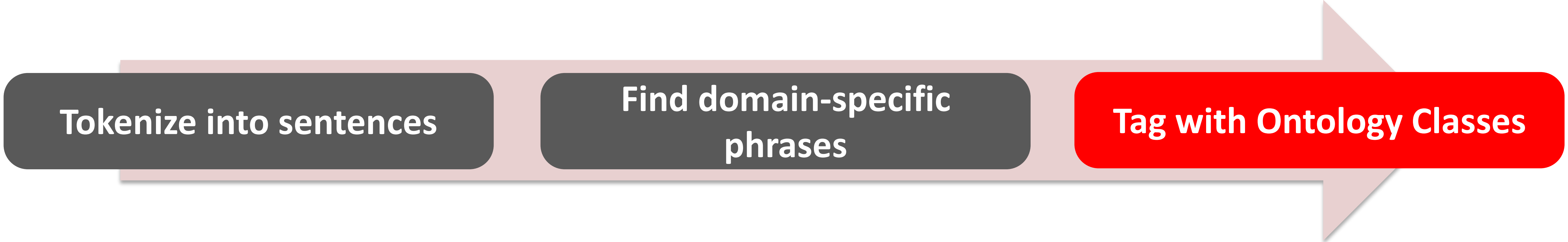
Tag with Ontology Classes

We have raised a request #9646604 and found that the **(device)/ReplaceableEntity** xxx-xxx-xxx-130a **(Power LED)/ReplaceableEntity** is **(amber)/Condition** and it is in **(hung state)/ProblemCondition**.

We **checked** the **device** for **connectivity issues**, **cleaned** the **fiber** and found that the **power supply unit** is **faulty**.

We **replaced** the **power supply unit**.

Putting it All Together (1/2): Tagging a Ticket



Tokenize into sentences

Find domain-specific phrases

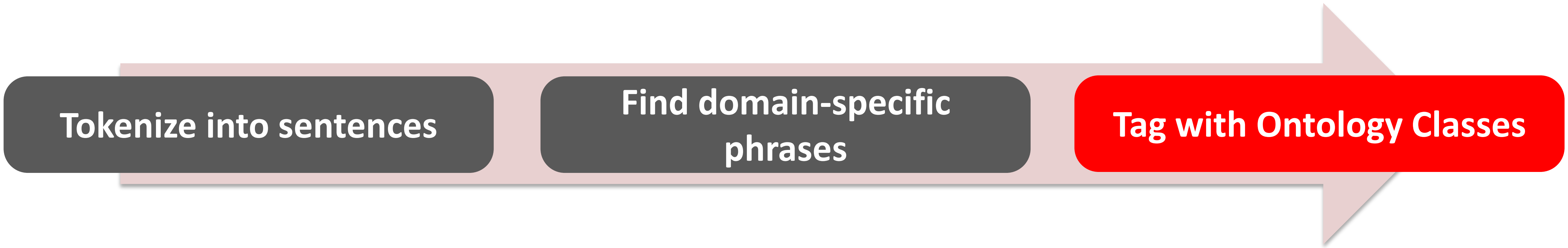
Tag with Ontology Classes

We have raised a request #9646604 and found that the (device)/ReplaceableEntity xxx-xxx-xxx-130a (Power LED)/ReplaceableEntity is (amber)/Condition and it is in (hung state)/ProblemCondition.

We (checked)/MaintenanceAction the (device)/ReplaceableEntity for (connectivity issues)/ProblemCondition, (cleaned)/MaintenanceAction the (fiber)/ReplaceableEntity and found that the (power supply unit)/ReplaceableEntity is (faulty)/ProblemCondition.

We **replaced** the **power supply unit**.

Putting it All Together (1/2): Tagging a Ticket



Tokenize into sentences

Find domain-specific phrases

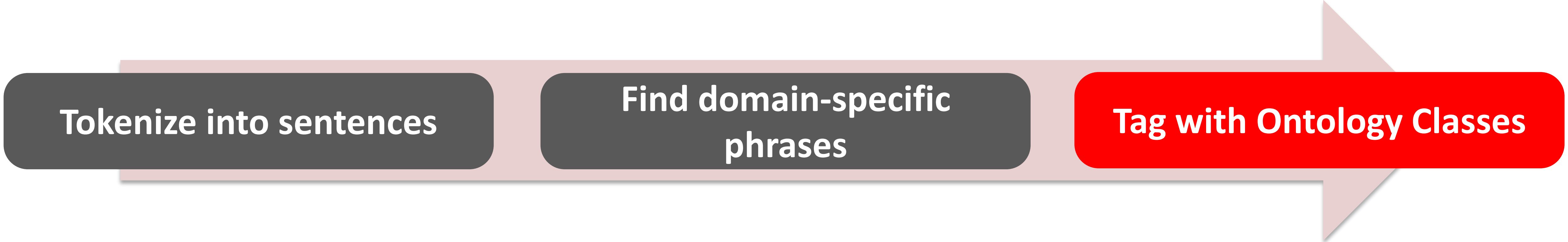
Tag with Ontology Classes

We have raised a request #9646604 and found that the (device)/ReplaceableEntity xxx-xxx-xxx-130a (Power LED)/ReplaceableEntity is (amber)/Condition and it is in (hung state)/ProblemCondition.

We (checked)/MaintenanceAction the (device)/ReplaceableEntity for (connectivity issues)/ProblemCondition, (cleaned)/MaintenanceAction the (fiber)/ReplaceableEntity and found that the (power supply unit)/ReplaceableEntity is (faulty)/ProblemCondition.

We (replaced)/PhysicalAction the (power supply unit)/ReplaceableEntity.

Putting it All Together (1/2): Tagging a Ticket



We have raised a request #9646604 and found that the **(device)/ReplaceableEntity** xxx-xxx-xxx-130a **(Power LED)/ReplaceableEntity** is **(amber)/Condition** and it is in **(hung state)/ProblemCondition**.

We **(checked)/MaintenanceAction** the **(device)/ReplaceableEntity** for **(connectivity issues) /ProblemCondition**, **(cleaned)/MaintenanceAction** the **(fiber)/ReplaceableEntity** and found that the **(power supply unit)/ReplaceableEntity** is **(faulty)/ProblemCondition**.

We **(replaced)/PhysicalAction** the **(power supply unit)/ReplaceableEntity**.

Putting it All Together (2/2): Information Inference

	Rule	Inference
Problems	**Entity** precedes/succeeds **ProblemCondition**	
Activities	**Entity|Condition** precedes/succeeds **MaintenanceAction**	
Actions	**Entity** precedes/succeeds **PhysicalAction**	

Putting it All Together (2/2): Information Inference

	Rule	Inference
Problems	**Entity** precedes/succeeds **ProblemCondition**	
Activities	**Entity|Condition** precedes/succeeds **MaintenanceAction**	
Actions	**Entity** precedes/succeeds **PhysicalAction**	

We have raised a request #9646604 and found that the **(device)/ReplaceableEntity** xxx-xxx-xxx-130a **(Power LED)/ReplaceableEntity** is **(amber)/Condition** and it is in **(hung state)/ProblemCondition**.

We **(checked)/MaintenanceAction** the **(device)/ReplaceableEntity** for **(connectivity issues)/ProblemCondition**, **(cleaned)/MaintenanceAction** the **(fiber)/ReplaceableEntity** and found that the **(power supply unit)/ReplaceableEntity** is **(faulty)/ProblemCondition**.

We **(replaced)/PhysicalAction** the **(power supply unit)/ReplaceableEntity**.

Putting it All Together (2/2): Information Inference

	Rule	Inference
Problems	**Entity** precedes/succeeds **ProblemCondition**	<device : hung state> <power supply unit : faulty>
Activities	**Entity|Condition** precedes/succeeds **MaintenanceAction**	
Actions	**Entity** precedes/succeeds **PhysicalAction**	

We have raised a request #9646604 and found that the **(device)/ReplaceableEntity** xxx-xxx-xxx-130a **(Power LED)/ReplaceableEntity** is **(amber)/Condition** and it is in **(hung state)/ProblemCondition**.

We **(checked)/MaintenanceAction** the **(device)/ReplaceableEntity** for **(connectivity issues)/ProblemCondition**, **(cleaned)/MaintenanceAction** the **(fiber)/ReplaceableEntity** and found that the **(power supply unit)/ReplaceableEntity** is **(faulty)/ProblemCondition**.

We **(replaced)/PhysicalAction** the **(power supply unit)/ReplaceableEntity**.

Putting it All Together (2/2): Information Inference

	Rule	Inference
Problems	**Entity** precedes/succeeds **ProblemCondition**	<device : hung state> <power supply unit : faulty>
Activities	**Entity|Condition** precedes/succeeds **MaintenanceAction**	<connectivity issues : checked> <fiber : cleaned>
Actions	**Entity** precedes/succeeds **PhysicalAction**	

We have raised a request #9646604 and found that the **(device)/ReplaceableEntity** xxx-xxx-xxx-130a **(Power LED)/ReplaceableEntity** is **(amber)/Condition** and it is in **(hung state)/ProblemCondition**.

We **(checked)/MaintenanceAction** the **(device)/ReplaceableEntity** for **(connectivity issues)/ProblemCondition**, **(cleaned)/MaintenanceAction** the **(fiber)/ReplaceableEntity** and found that the **(power supply unit)/ReplaceableEntity** is **(faulty)/ProblemCondition**.

We **(replaced)/PhysicalAction** the **(power supply unit)/ReplaceableEntity**.

Putting it All Together (2/2): Information Inference

	Rule	Inference
Problems	**Entity** precedes/succeeds **ProblemCondition**	<device : hung state> <power supply unit : faulty>
Activities	**Entity|Condition** precedes/succeeds **MaintenanceAction**	<connectivity issues : checked> <fiber : cleaned>
Actions	**Entity** precedes/succeeds **PhysicalAction**	<power supply unit : replace>

We have raised a request #9646604 and found that the **(device)/ReplaceableEntity** xxx-xxx-xxx-130a **(Power LED)/ReplaceableEntity** is **(amber)/Condition** and it is in **(hung state)/ProblemCondition**.

We **(checked)/MaintenanceAction** the **(device)/ReplaceableEntity** for **(connectivity issues) /ProblemCondition**, **(cleaned)/MaintenanceAction** the **(fiber)/ReplaceableEntity** and found that the **(power supply unit)/ReplaceableEntity** is **(faulty)/ProblemCondition**.

We **(replaced)/PhysicalAction** the **(power supply unit)/ReplaceableEntity**.

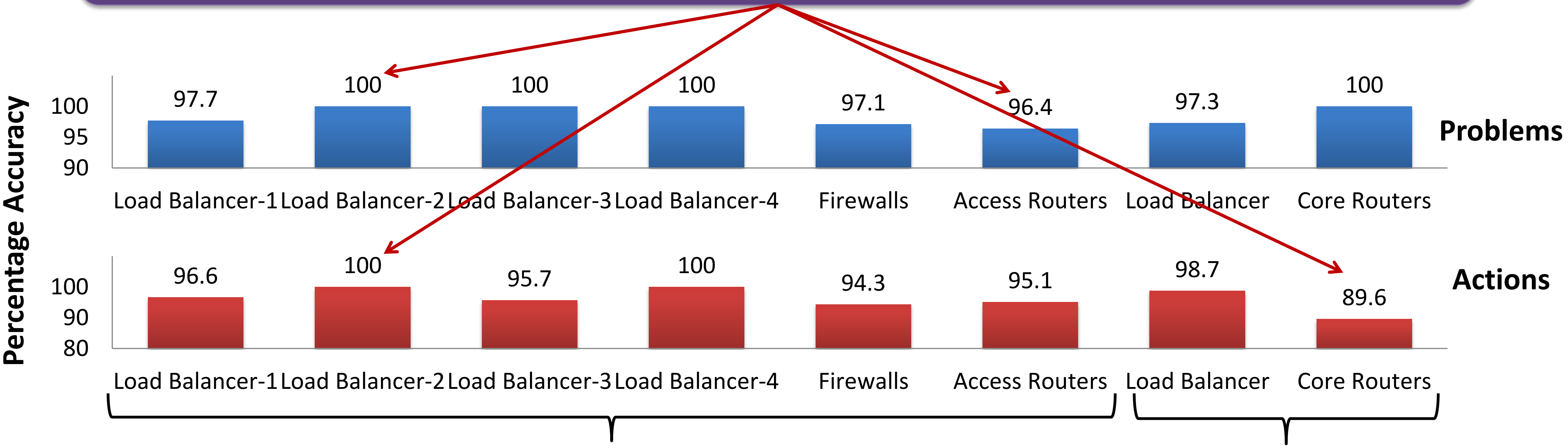
NetSieve Evaluation

Evaluation Methodology

- **Goals: Evaluate Accuracy and Usability**
  - **Metrics:**
    - Percentage Accuracy, F-Score, Precision, Recall
    - Time to read a ticket manually vs. NetSieve inference
  - **Dataset:** 10K+ tickets
    - Ground truth: 696 tickets labeled by an expert; 155 tickets from two network vendors
  - **Method:**
    1. Compare expert-labeled Problems and Actions with NetSieve inference
    2. Survey of five operators each shown 20 tickets at random

Evaluating Accuracy: Expert-labeled and Vendor Tickets

96%-100% accuracy for Problems; 89%-100% accuracy for Actions



Expert-labeled (696 tickets)

Vendor (155 tickets)

NetSieve Use Cases for Network Management

	Team	Questions	Findings
1	**Capacity Planning**	Why is network redundancy ineffective?	1. Faulty cables 2. Software version mismatch 3. Misconfigurations
2	**Incident Management**	What are the top-k failing components?	1. Line card failures 2. Defective memory 3. Supervisor engine
3	**Network Architecture**	Are new devices more reliable?	1. A new access router is half as reliable as its predecessor 2. Software bugs dominated failures in one type of load balancers

Conclusion

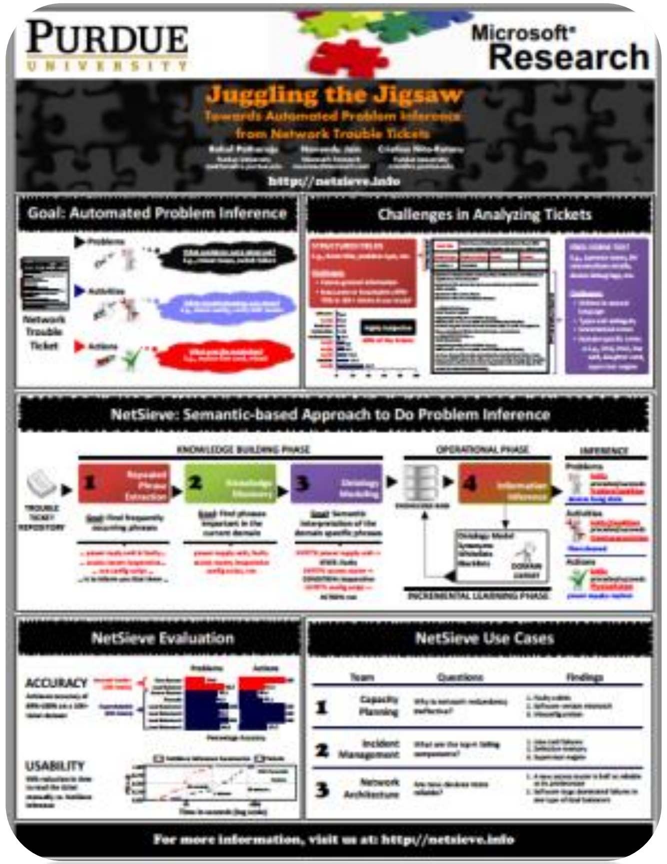
- Goal: Automate problem inference from trouble tickets

- NetSieve semantic based approach
  - Combines NLP, knowledge discovery and ontology modeling in a novel way
  - Three key features: Problems, Activities and Actions
  - Achieves an accuracy of 83%-100% over a large ticket dataset

- Future Work
  - Build an ontology model automatically
  - Improving accuracy using expert feedback
  - Applying NetSieve to other problem domains

Poster & Demo session

Tomorrow Evening!



PURDUE UNIVERSITY

Microsoft® Research

Juggling the Jigsaw

Towards Automated Problem Inference from Network Trouble Tickets

Rahul Potharaju, Navendu Jain, Cristina Nita-Rotaru

http://netsieve.info

Goal: Automated Problem Inference

Challenges in Analyzing Tickets

NetSieve: Semantic-based Approach to Do Problem Inference

NetSieve Evaluation

NetSieve Use Cases

	Team	Questions	Findings
1	Capacity Planning	Why is network redundancy ineffective?	[illegible]
2	Incident Management	What are the top-k failing components?	[illegible]
3	Network Architecture	Are new devices more reliable?	[illegible]

For more information, visit us at: http://netsieve.info

Thank You
for your time!

Project page

http://netsieve.info

NetSieve Problem Inference

Towards an automated problem inference system.

Part of a larger effort, NetWiser, that aims to improve service reliability in datacenters.

See what NetSieve can do for you...

Systematic Workflow

Failure Trends

Problem Inference

Ontology Model

Architectural Overview

Following is the architectural view of our system.