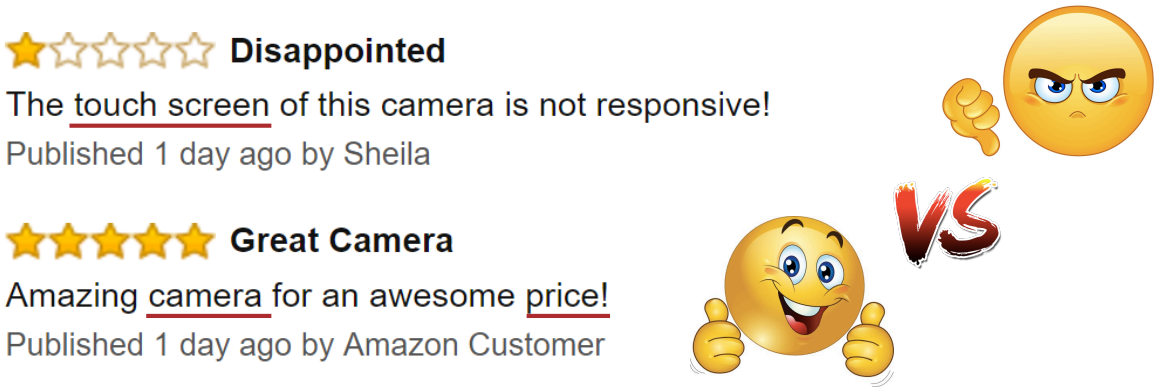
Lifelong-RL: Lifelong Relaxation Labeling for Separating Entities and Aspects in Opinion Targets

Lei Shu, University of Illinois at Chicago, USA
Bing Liu, University of Illinois at Chicago, USA
Hu Xu, University of Illinois at Chicago, USA
Annice Kim, RTI International, USA
EMNLP, 2016

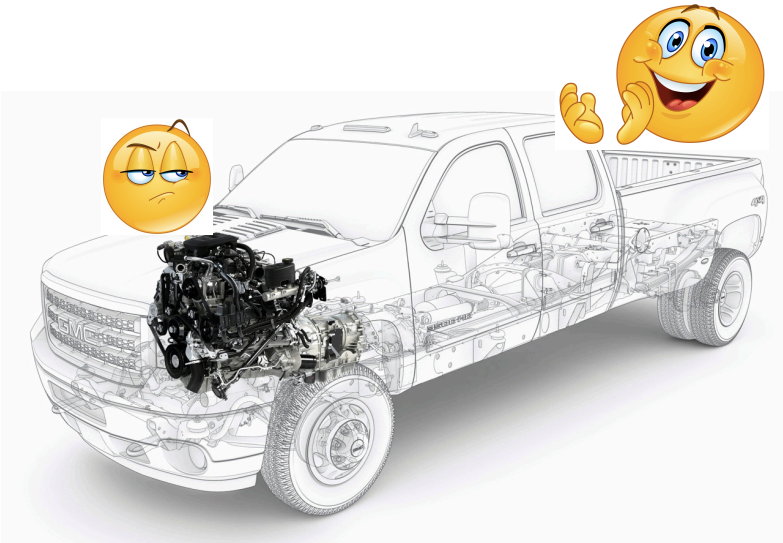
Sentiment Analysis on Customer Reviews
(Liu, 2012)

- Sentiment analysis:
  - identify opinions (sentiment polarities) from a piece of text;
  - multiple levels of sentiment analysis: document-level, sentence-level, aspect-level sentiment analysis.
- Aspect-level sentiment analysis on customer reviews:
  - identify **opinion targets (targets for short)** and classify the opinions on those targets.



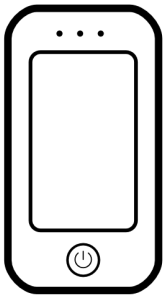
Opinion Target

- An opinion target can be
  - an **entity** (general aspect):
    - I like this **car**.
  - or an **aspect** (part or attribute) of an entity:
    - The **engine** of this car is great.
- Why to separate entities and aspects?
  - Opinion on an entity is an opinion as a **whole**;
  - Opinion on an aspect is just for that aspect.
  - Example:
    - although the **engine** is lightly weak, this **car** is great.



What is an Entity ?

- In this paper, an entity can be:
  - named entity (“Apple”, “iPhone”);
  - product category (“Phone”);
  - abstract product (“machine”, “product”).



What is an Aspect ?

- In this paper, an aspect can be:
  - a part;
  - or an attribute of an entity.

- Examples:
  - The camera of this phone is great.
  - The phone's price is great.

Our Goal

- Assign each target a **target label** as entity, aspect or NIL in an unsupervised manner.

- Aspects are shared in many products, we can utilize lifelong machine learning to improve the classification.

- Note that we only focus on classifying entity and aspect after the opinion targets have been extracted.

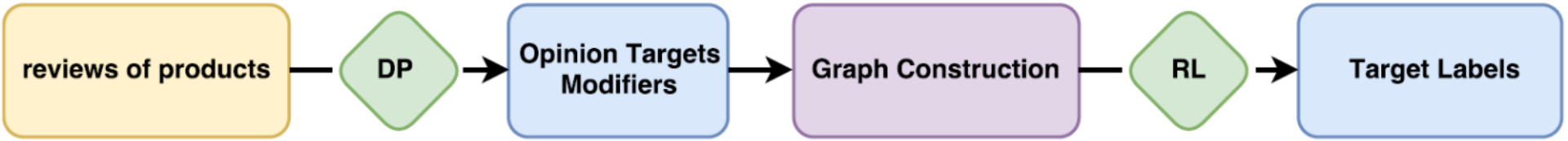
The picture is from AAAI 2013 Spring Symposium

Roadmap

- **Framework Overview**
- Relaxation Labeling
- Lifelong Relaxation Labeling
- Experiment
- Conclusion

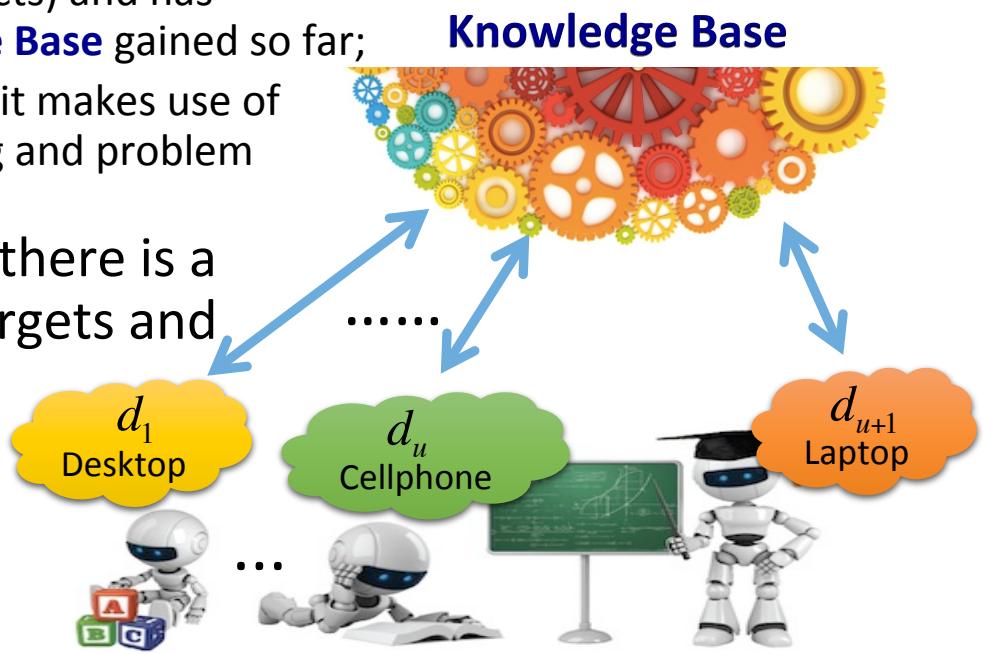
Labeling Pipeline

- We use an unsupervised opinion target extraction method **Double Propagation (DP)** (Qiu et. al, 2011) to get a list of targets;
- We use targets and modifiers (text clues) from reviews to perform graph construction and use **Relaxation Labeling (RL)** (Hummel et. al, 1986) to classify targets as entity, aspect or NIL.

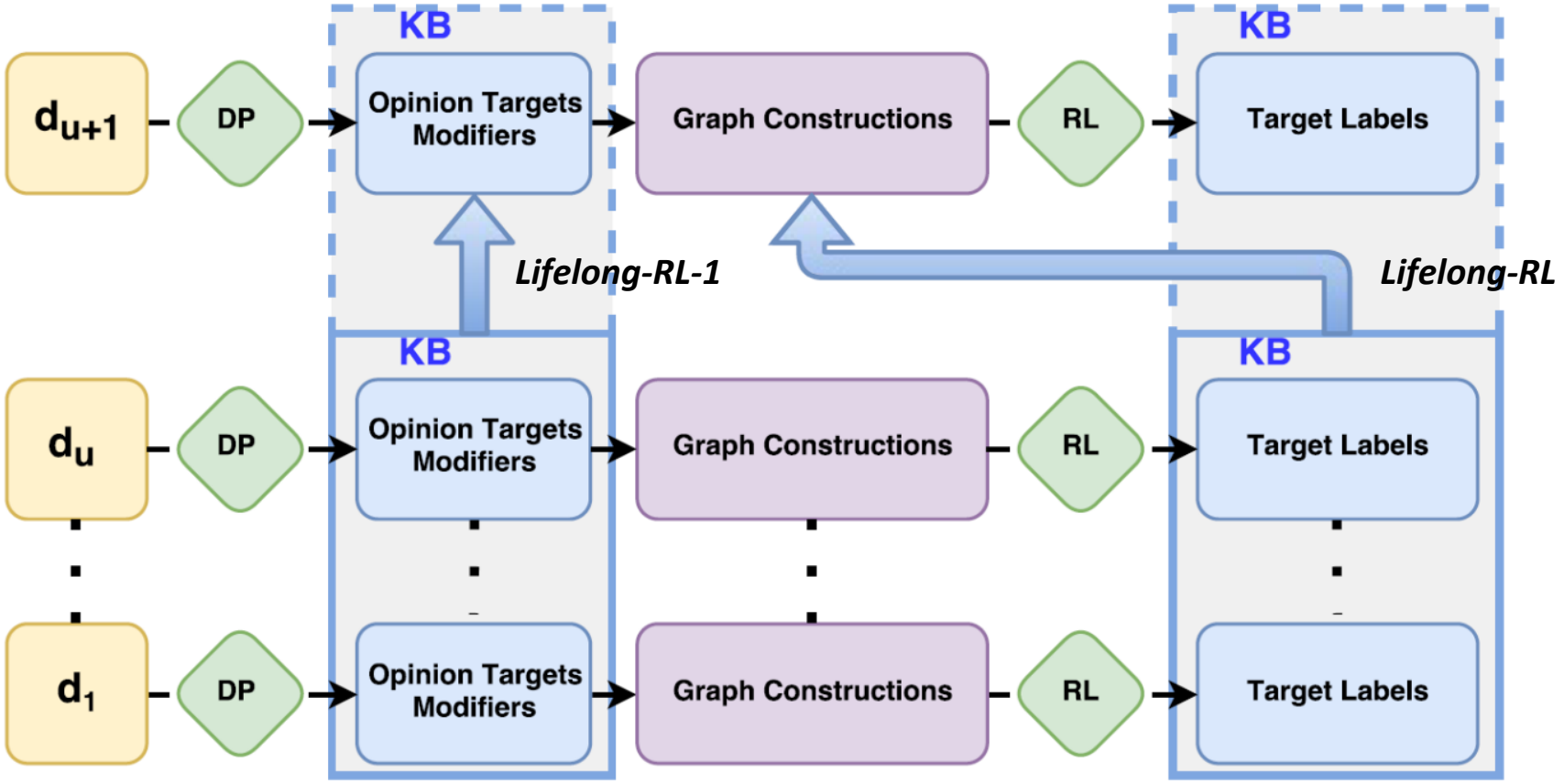


Lifelong Machine Learning (LML) (Thrun, 1996; Silver et. al, 2013; Chen and Liu, 2014)

- LML works as how human learn:
  - Assume the learner has performed a number of learning tasks $d_1 \ldots d_u$ in the past domains (products) and has retained the knowledge in a **Knowledge Base** gained so far;
  - In the new/current task (product) d_{u+1} , it makes use of past knowledge to help current learning and problem solving.
- The approach is effective because there is a significant amount of sharing of targets and target relations across products.



Lifelong Relaxation Labeling Framework

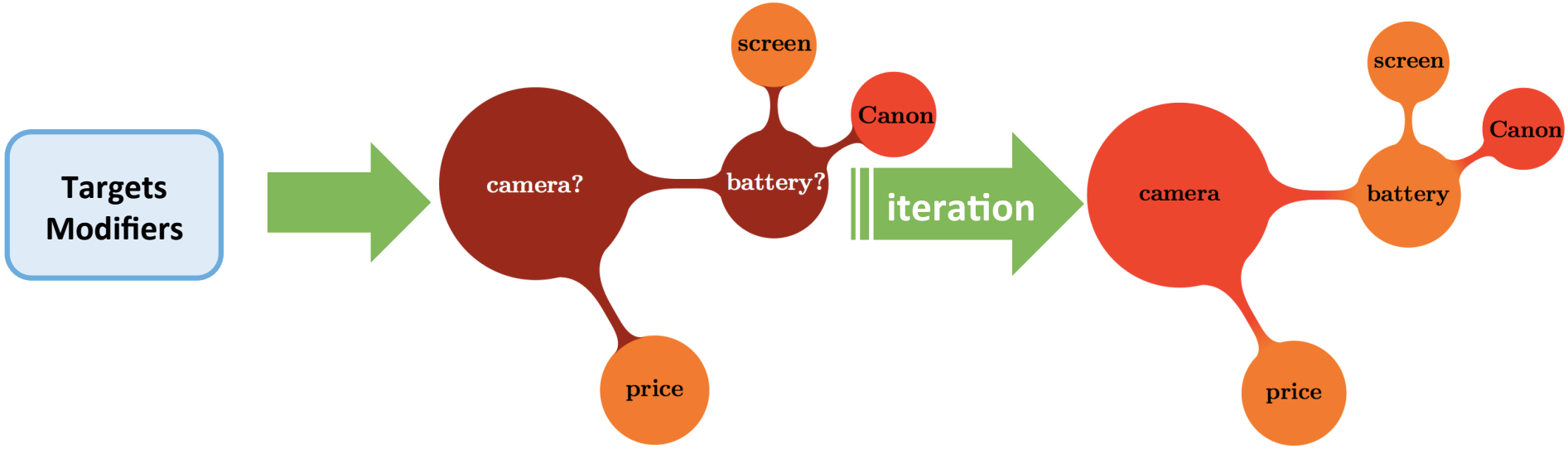


Roadmap

- Framework Overview
- **Relaxation Labeling**
- Lifelong Relaxation Labeling
- Experiment
- Conclusion

Relaxation Labeling (RL)

- An unsupervised graph-based label propagation algorithm;
- A node is a target and an edge indicates how two targets influence each other;
- The graph is constructed from targets and modifiers.



Modifiers (Text Clues) as Features

- Use **type modifiers** to determine the initial label distribution of nodes:
  - entity modifiers: “this”, “these” e.g., “**this** camera is great” indicates that “camera” is probably an entity;
  - aspect modifiers: implicitly assumed when the appearance of entity modifiers is inadequate.
- Use **relation modifiers** to determine the structure (existence of an edge) of the graph:
  - **conjunction modifier**: “price **and** service” (if price is an aspect service is an aspect and vice versa);
  - **entity-aspect modifier**: “camera **’s** price”;
  - **aspect-entity modifier**: “camera **’s** price”.

Initial Distribution for Nodes and Edges

- Each node in the graph is associated with a distribution on labels {entity, aspect, NIL} that is updated during iteration;
- $P(L(t{\downarrow}i))=[\blacksquare\cdots@\cdots@\cdots]$
- Each edge is associated with a fixed 3x3 matrix representing conditional probabilities of how the label distribution of one target affects the other's.
- $P(L(t{\downarrow}i)|L(t{\downarrow}j))=[\blacksquare\cdots\&\cdots\&\cdots@\cdots\&\cdots\&\cdots@\cdots\&\cdots\&\cdots]$
- Distribution on labels are iteratively updated until convergence.

Roadmap

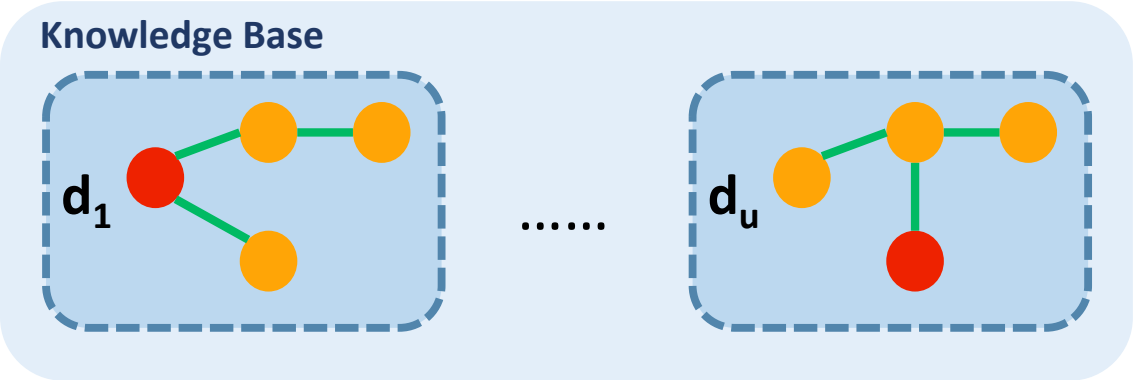
- Framework Overview
- Relaxation Labeling
- **Lifelong Relaxation Labeling**
- Experiment
- Conclusion

Lifelong Relaxation Labeling (Lifelong-RL)

- RL on a single domain data may not be sufficient
  - **Insufficient edges**: relation modifiers for constructing edges;
  - **Unreliable initial probabilities**: type modifiers for setting up ''seed'' nodes.
- Lifelong learning comes to rescue:
  - Assume the learner has "relax labeled" u domains (products);
  - The learner can borrow some useful prior/past knowledge (e.g., nodes and edges etc.) in the Knowledge Base (KB) to help RL in the new/current domain d_{u+1}
  - The learner can further add the results of the new/current domain d_{u+1} to the KB for future use.

Knowledge Base (KB)

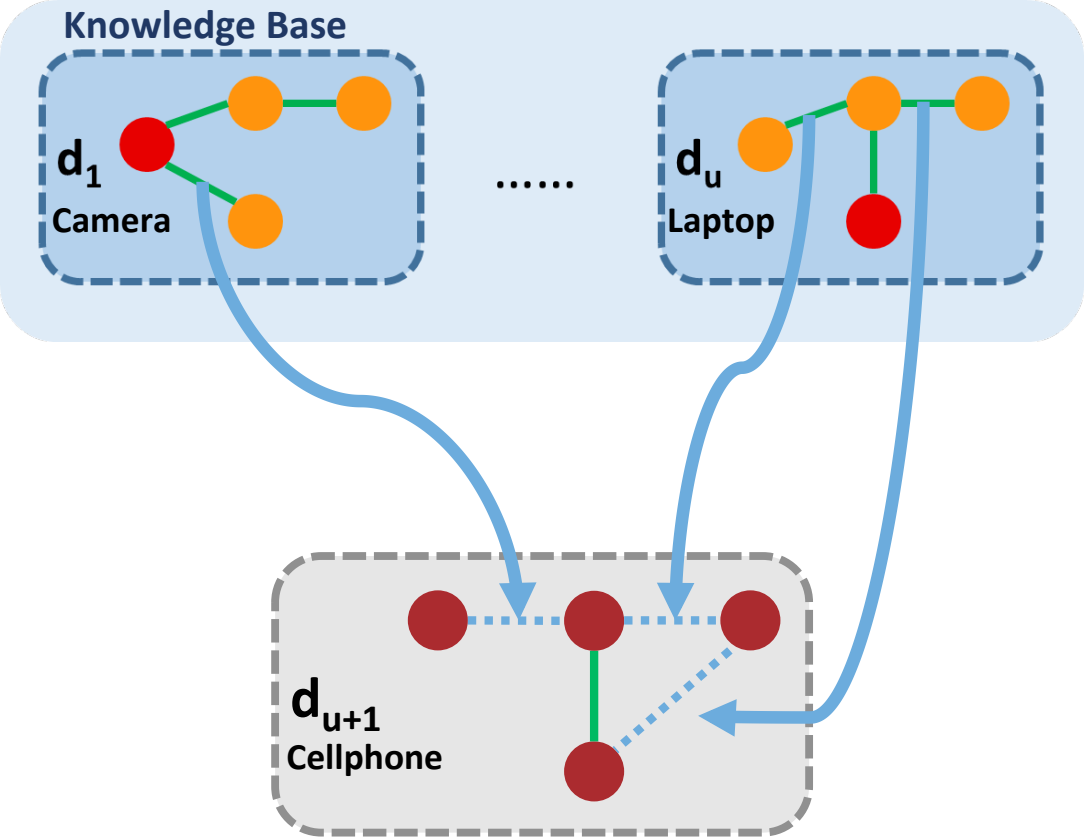
- The results of RL for each domain is stored in a KB
- There are two types of knowledge:
  - prior edges:
    - relation modifiers;
    - used in **Lifelong-RL-1** and **Lifelong-RL**;
  - prior labels:
    - target labels;
    - used in **Lifelong-RL**.



Lifelong-RL-1: Exploiting Relation Modifiers

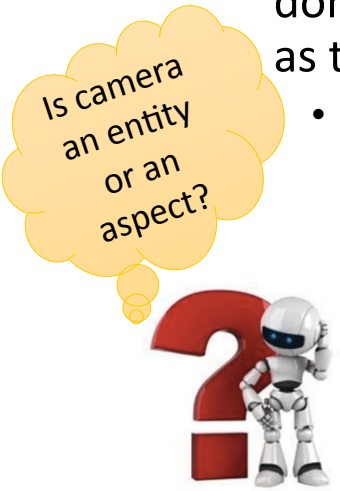
- Idea: relation modifiers from past domains can help to link targets in the current domain.

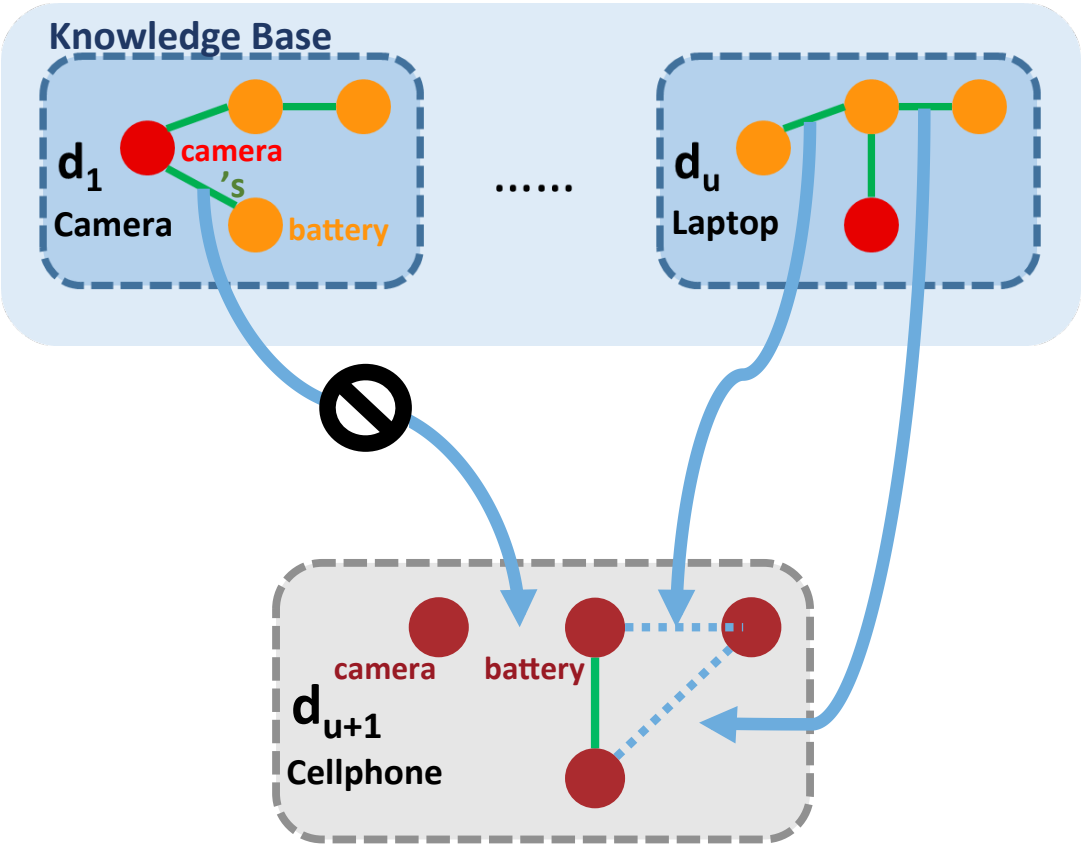
- Example:
  - Relation modifiers can be shared across domains;
  - Example: a Cellphone domain can borrow the edge "camera and battery" from Laptop.



Consistency Check

- Knowledge may contain noise:
  - (1) target labels may be wrong since Double Propagation and RL may be inaccurate;
  - (2) target labels in one domain may not be the same as the other;
    - Example: a new/current Cellphone domain cannot borrow the relation modifier "camera 's battery" from Camera domain.





Consistency Check

- Lifelong learner further performs a **consistency check**.
- Two types of consistency check:
  - **Label Consistency Check** (for noise (1)): ensure that a relation modifier from a domain in the KB are consistent with target labels of those two targets associated with that relation modifier in that past domain;
  - **Type Consistency Check** (for noise (2)): the type modifier for a target in new/current domain must match the type modifiers of that target in the past domains.

Lifelong-RL: Further Exploiting Target Labels

- Idea: target labels from past domains can give a better idea about the initial label distributions of targets in the current domain.
- Example:
  - after labeling domains like Laptop, Tablet and E-reader, it is likely that "camera" is an aspect in the new domain.

Roadmap

- Framework Overview
- Relaxation Labeling
- Lifelong Relaxation Labeling
- **Experiment**
- Conclusion

Experiment: Dataset & Evaluation

- Datasets:
  - Test Datasets: reviews on 8 products labeled with entities and aspects;
  - Unlabeled review datasets for LML: 100 diverse domains (products) (Chen and Liu 2014); Each domain has 1000 reviews.

- Evaluation Method:
  - precision, recall, F1-score and accuracy as evaluation measures.
  - We only evaluate entities and aspects.

Experiment: Compared Methods

- NER+TM
  - Use Named Entity Recognition (NER) (UIUC NER) to find entities;
  - Use type modifier(TM) to recognize more product category or abstract product as entities;
  - Regard those entities as predicted entities and the rest of the targets as aspects.
- NER+TM+DICT
  - Run NER+TM on the 100 datasets for LML to get a list of entities;
  - For a new task, treat targets in the list as an entity; the rest are aspects.
- RL
  - the base method that performs RL without the help of LML.
- Lifelong-RL-1
  - perform LML with RL but the current task only uses the prior edges in the KB.
- Lifelong-RL
  - improve Lifelong-RL-1 by further incorporating prior labels in the KB.

Results

Algorithm	Entity F1-score	Aspect F1-score
NER+TM	61.3	78
RL	61.2	77.2
Lifelong-RL-1	**77.4**	**79.7**
Lifelong-RL	**79.9**	**80.2**

For entity label, Lifelong-RL improves RL 23.3% on Recall 10.9% on Precision.

Roadmap

- Overview
- Relaxation Labeling
- Lifelong Relaxation Labeling
- Experiment
- **Conclusion**

Conclusion

- Propose the problem of classifying opinion targets into entities and aspects for sentiment analysis on reviews;
- Propose a novel method based on relaxation labeling and the paradigm of lifelong machine learning to solve the problem.
- Experimental results show the effectiveness of the proposed method and lifelong machine learning.

Question and Answer [?]

Lei Shu, Bing Liu, Hu Xu, and Annice Kim.
2016.
Lifelong-rl: Lifelong relaxation labeling for separating entities and aspects in opinion targets.
In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 225–235, Austin, Texas, November. Association for Computational Linguistics.