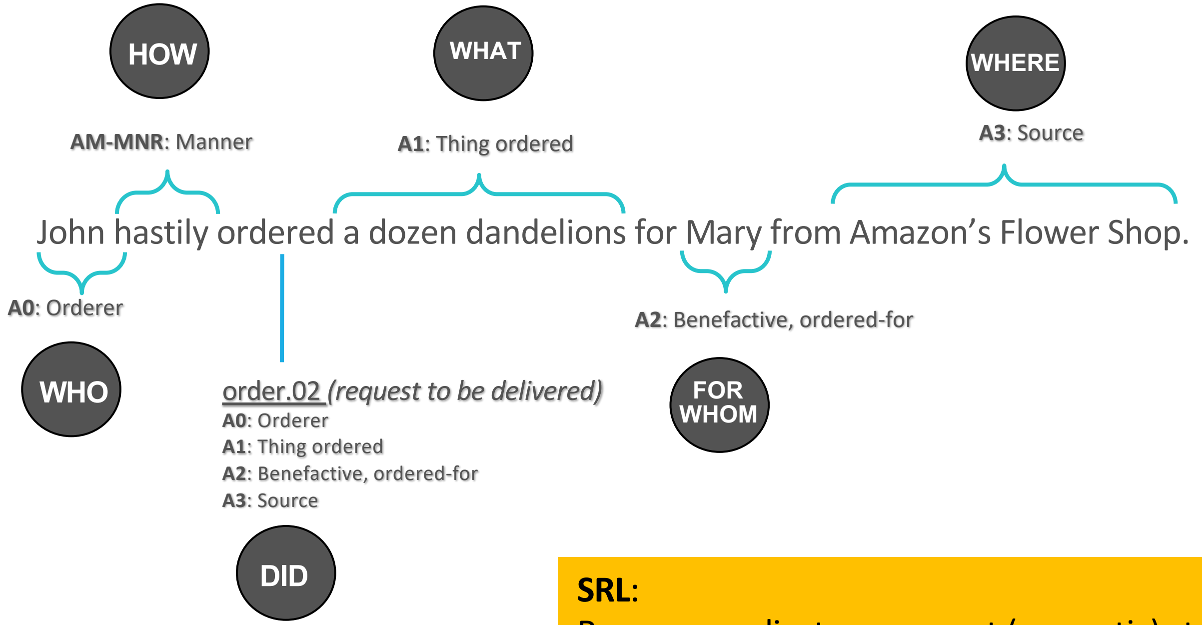
# Active Learning for Black-Box Semantic Role Labeling with Neural Factors

**Chenguang Wang, Laura Chiticariu, Yunyao Li**
IBM Research – Almaden

# Semantic Role Labeling (SRL)




**SRL**:
Recover predicate-argument (semantic) structure from text
Who did what to whom, when, where and how?

# Semantic Role Labeling (SRL)

HOW

**AM-MNR**: Manner

WHAT

**A1**: Thing ordered

WHERE

**A3**: Source

John hastily ordered a dozen dandelions for Mary from Amazon's Flower Shop.

**A0**: Orderer

WHO

**A2**: Benefactive, ordered-for

FOR WHOM

order.02 *(request to be delivered)*
**A0**: Orderer
**A1**: Thing ordered
**A2**: Benefactive, ordered-for
**A3**: Source

DID

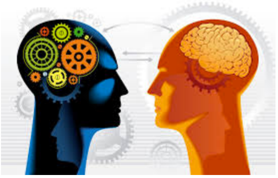
**SRL**:
Recover predicate-argument (semantic) structure from text
Who did what to whom, when, where and how?

General Applications


Information Extraction

WIKIPEDIA
The Free Encyclopedia

Knowledge Base Construction



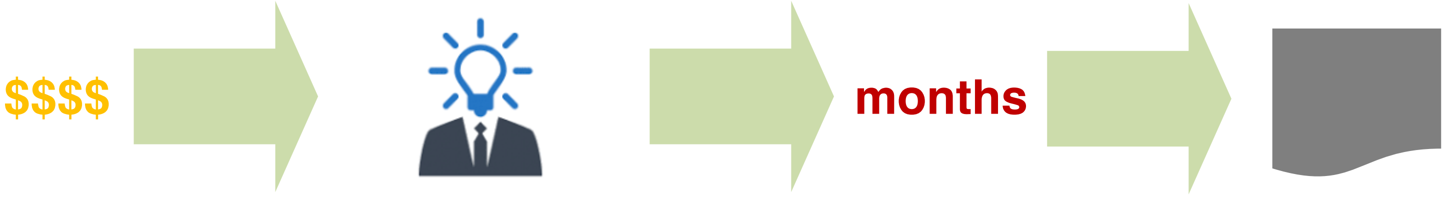
Question Answering

# Main Challenge Behind SRL

- **Key to improving SRL**
  - More SRL Labeled data

# Main Challenge Behind SRL

- **Key to improving SRL**
  - More SRL Labeled data
- **Way to generate SRL labeled data**




  - Label by experts
    - Issue: cost is high
  - Examples
    - PropBank, FrameNet

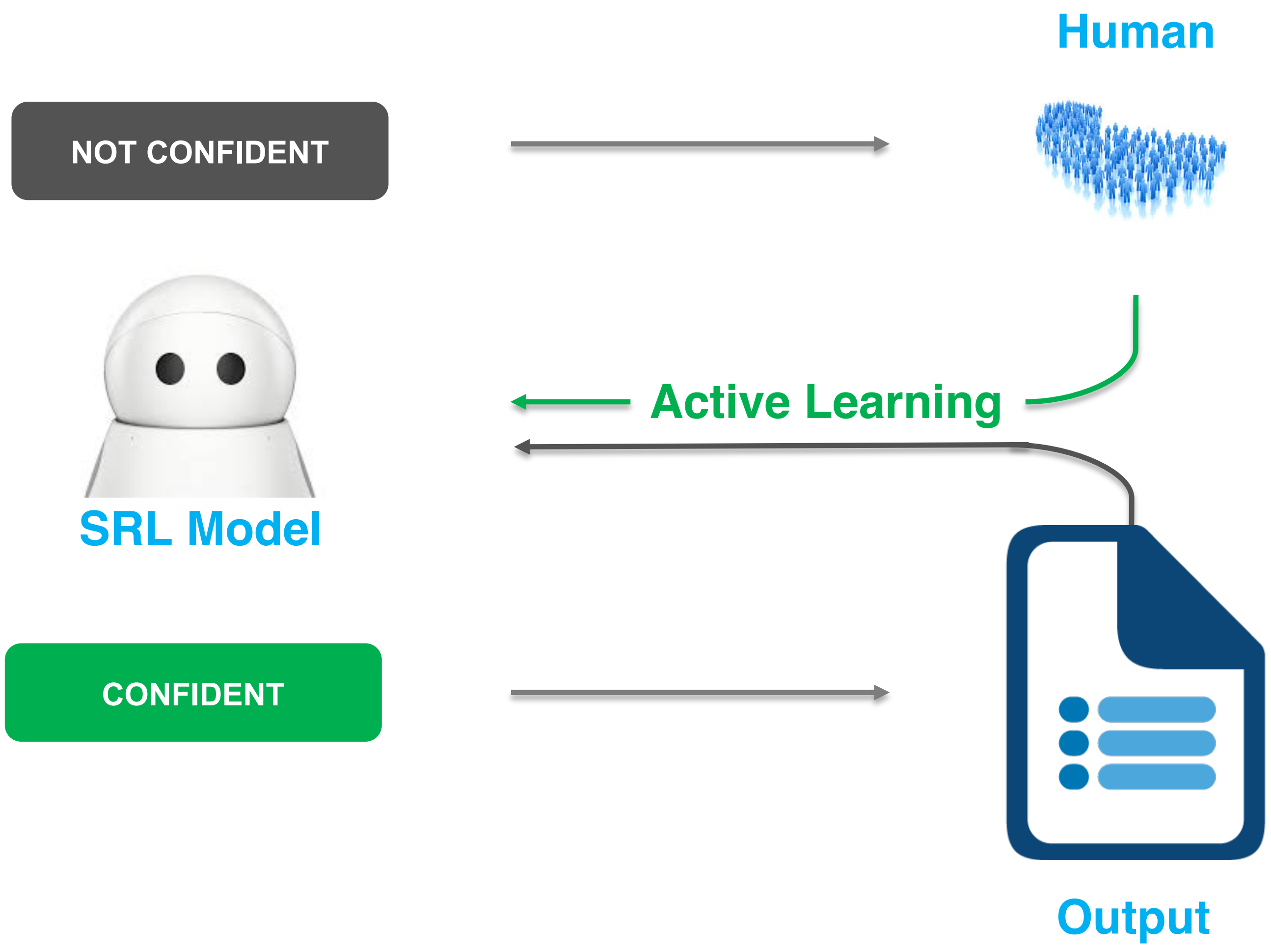
# Active Learning is Attractive for SRL

- Goal of Active Learning
  - carefully select the training data based on **query strategy** from which the model is being learnt in order to **achieve good performance with less training data**

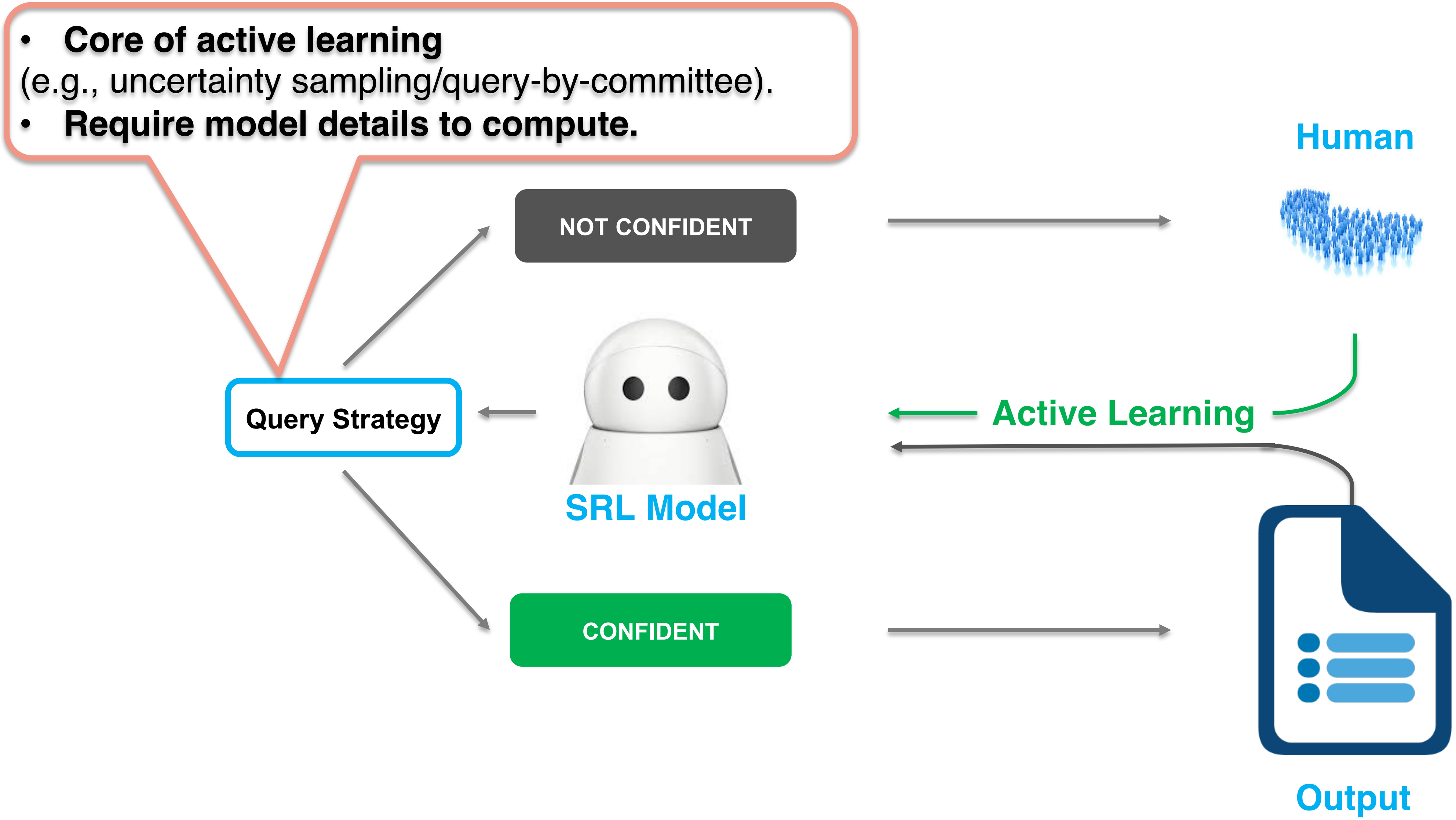
- Goal of Active Learning
  - carefully select the training data based on **query strategy** from which the model is being learnt in order to **achieve good performance with less training data**

How about directly apply **Active Learning** to **SRL**?

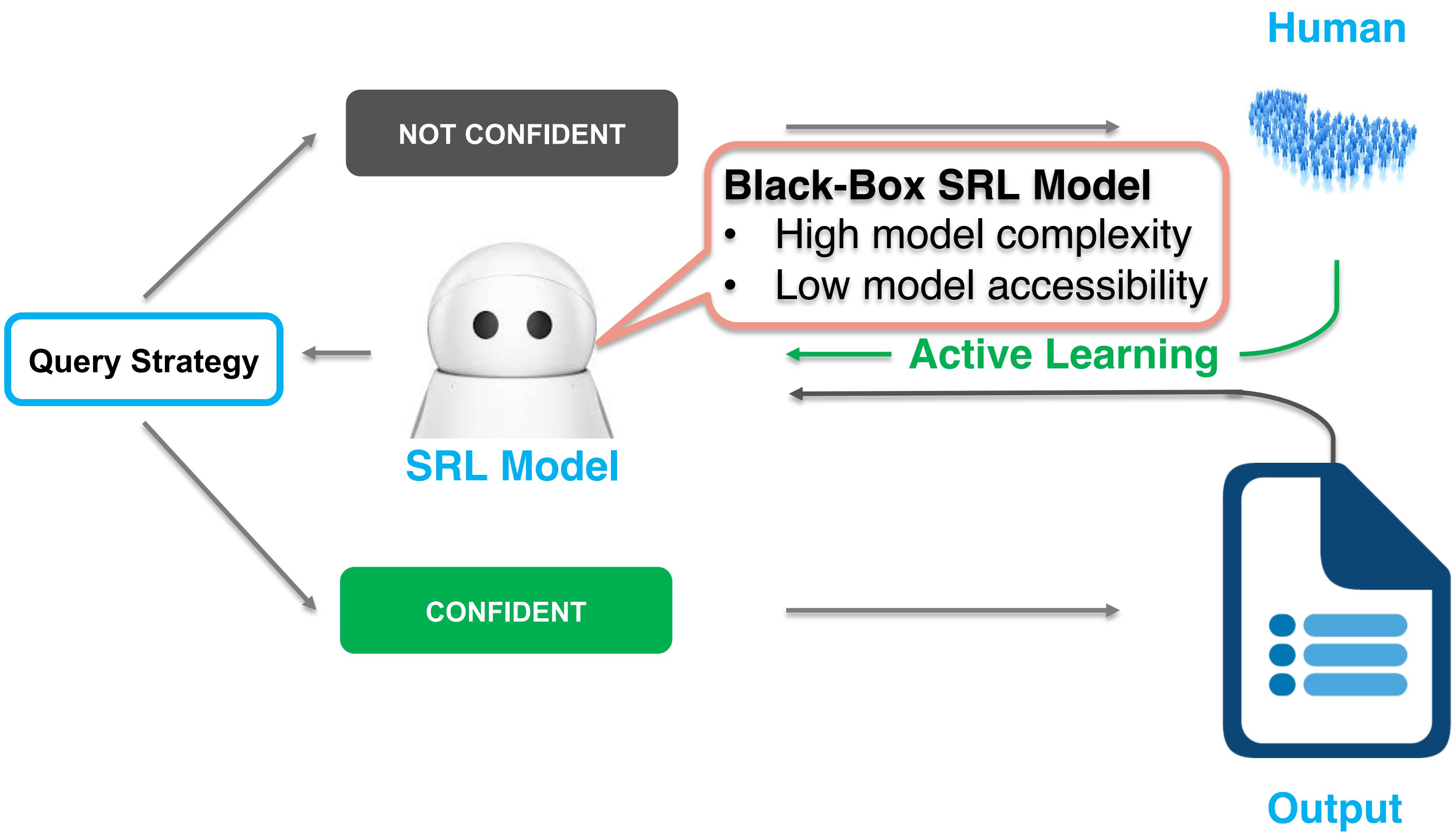
# Traditional Active Learning for SRL



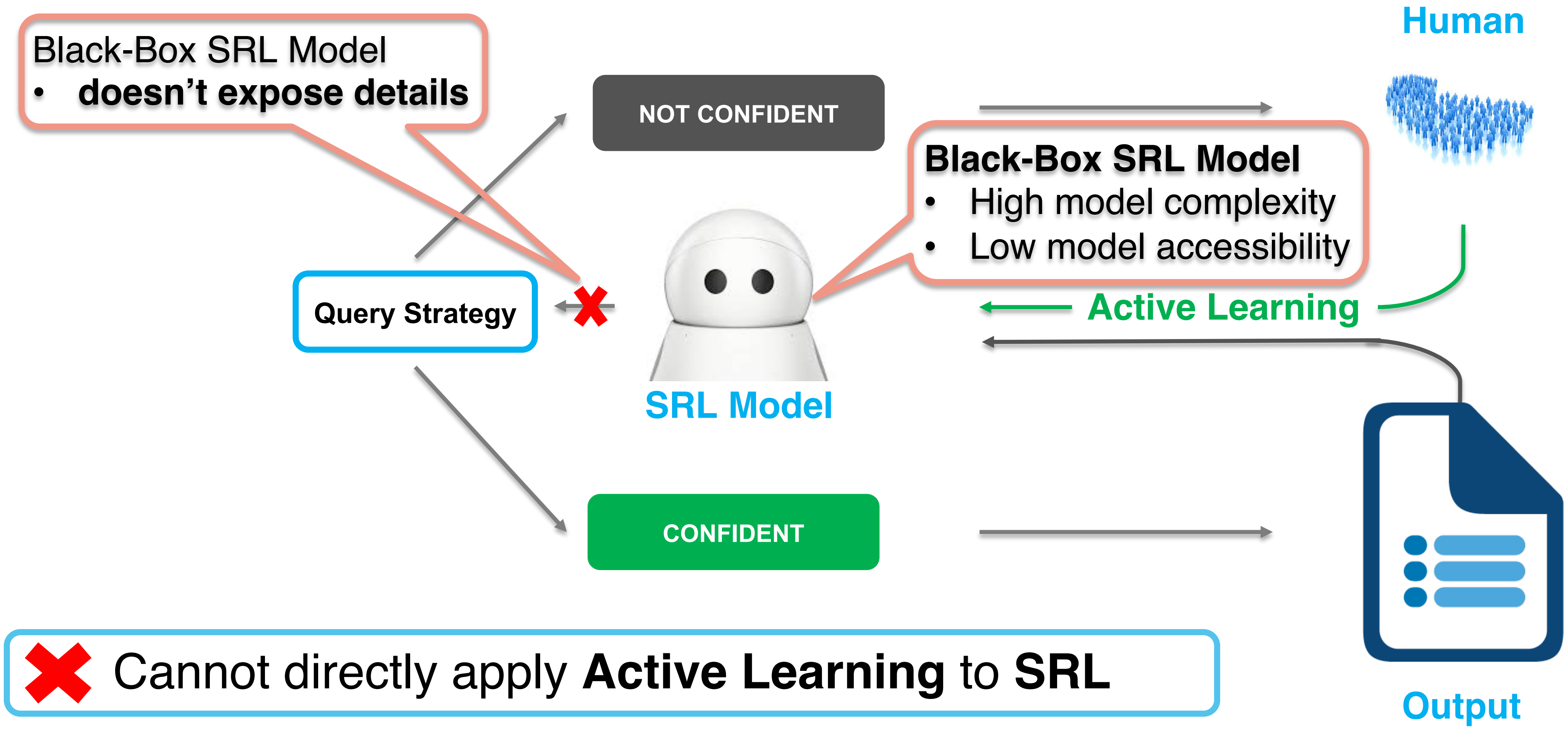
# Traditional Active Learning for SRL



IBM

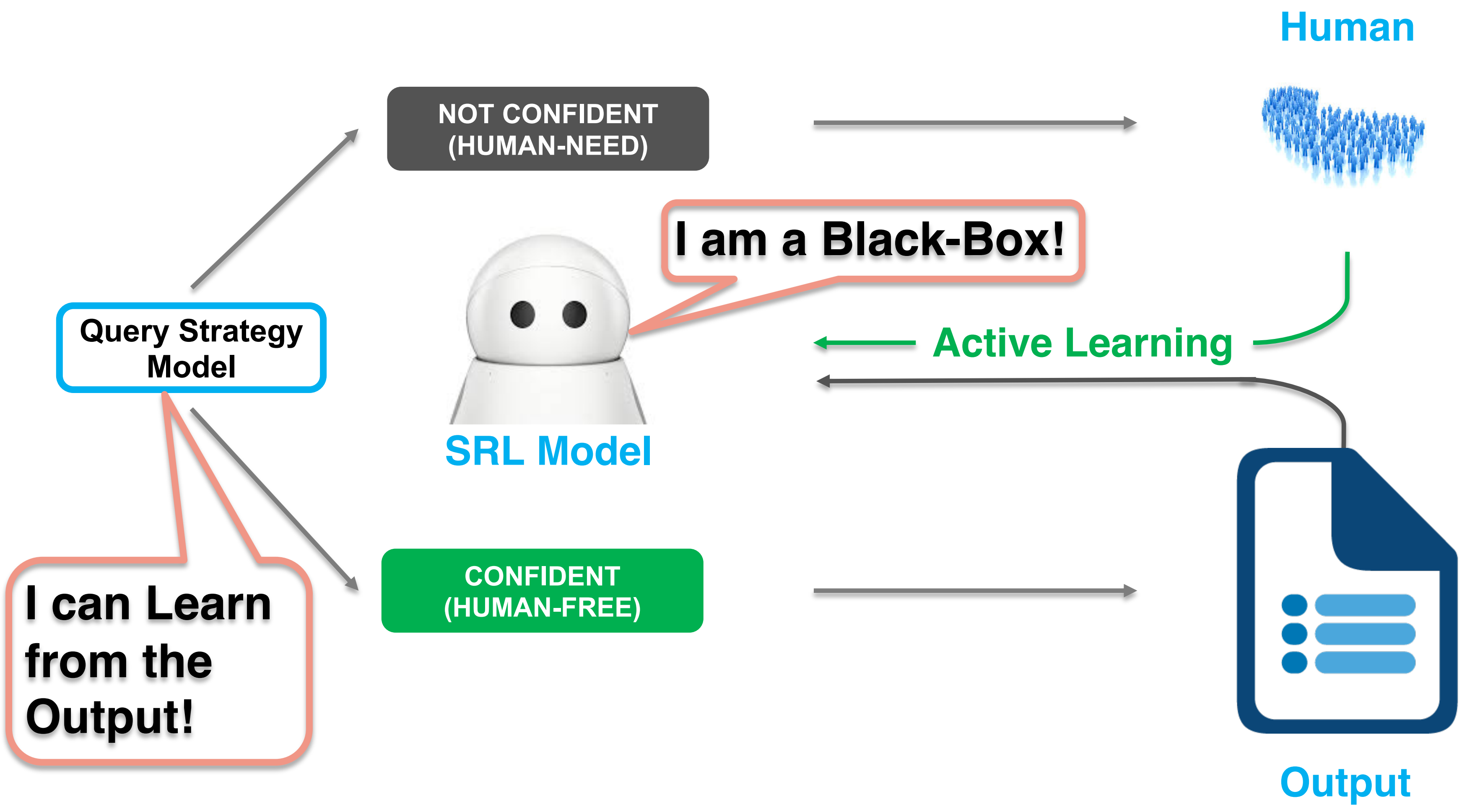

Human
NOT CONFIDENT
Black-Box SRL Model
• High model complexity
• Low model accessibility
Query Strategy
Active Learning
SRL Model
CONFIDENT
Output

# Traditional Active Learning for SRL




Cannot directly apply **Active Learning** to **SRL**

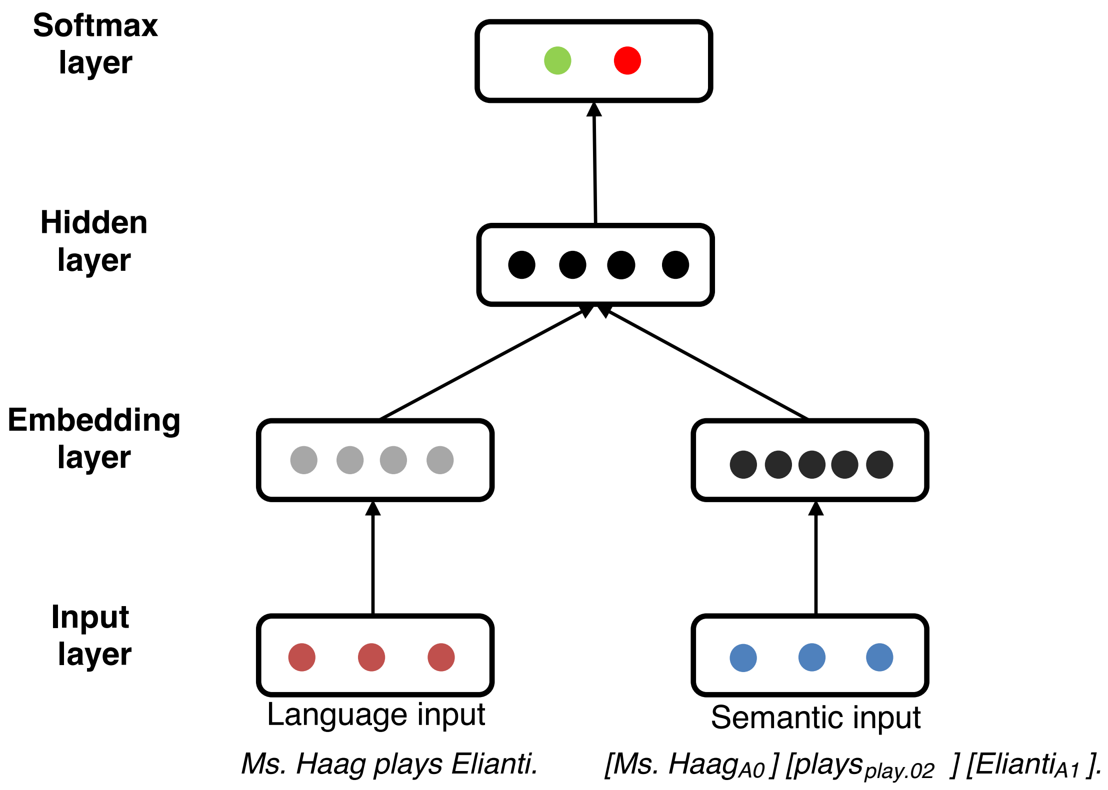
# **Our Approach:** Active Learning for Black-Box SRL



# **Core in Our Approach:** Neural Query Strategy Model

**Classification Model:** Classify a predicted SRL label based on the model output

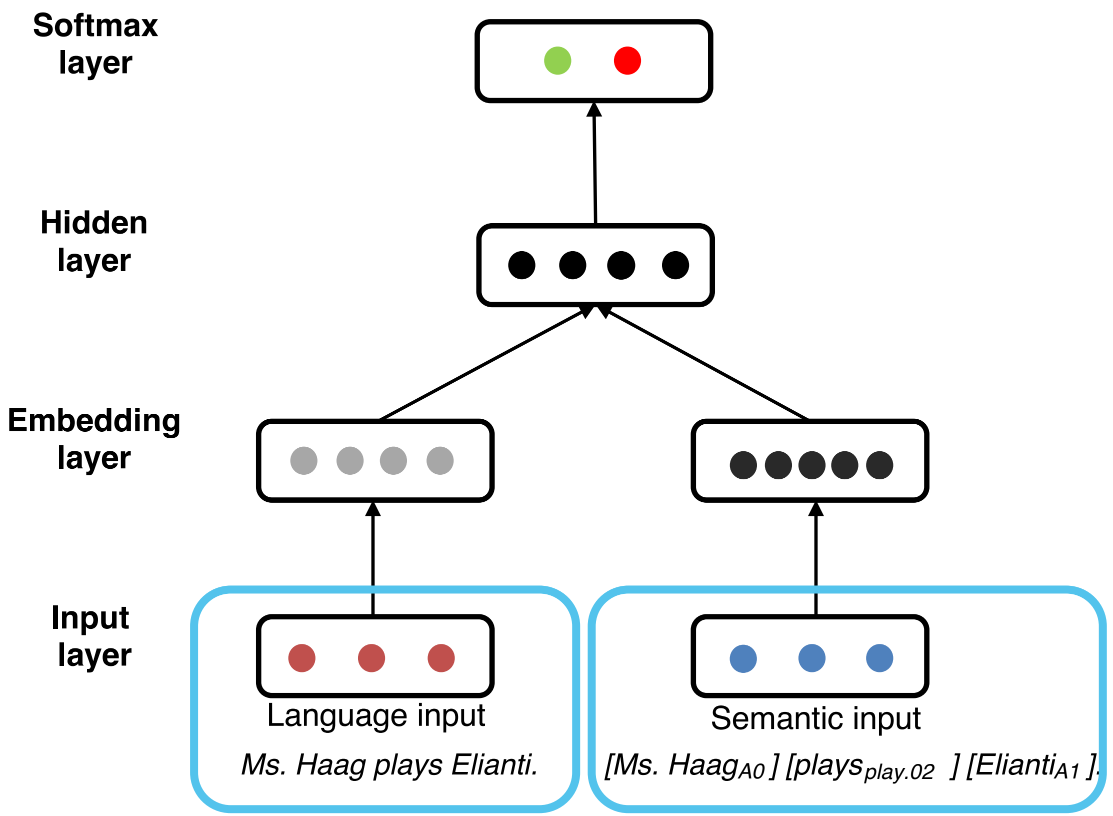
- **Human-free SRL** label if the predicted SRL label is likely to be the gold SRL label
- **Human-need SRL** label otherwise



# **Core in Our Approach:** Neural Query Strategy Model

**Input Layer**

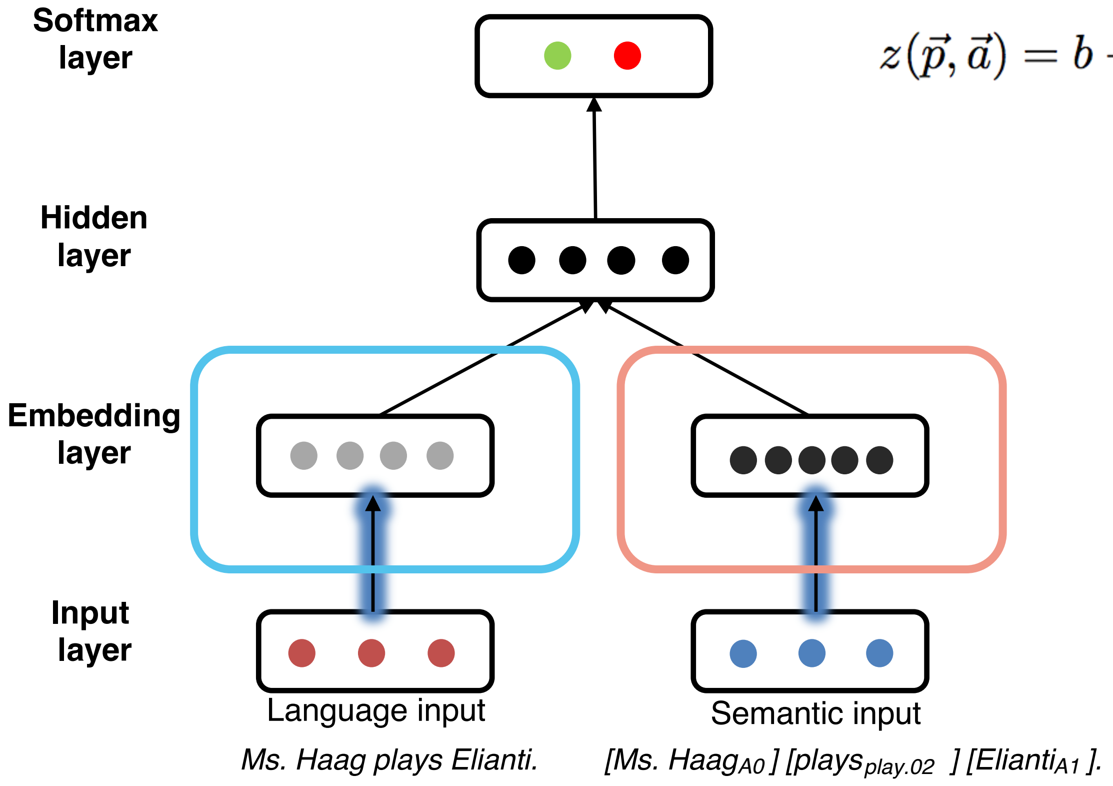
- **Language input:** Text/sentence
- **Semantic input:** Argument/role label, predicate/frame label



# **Core in Our Approach:** Neural Query Strategy Model

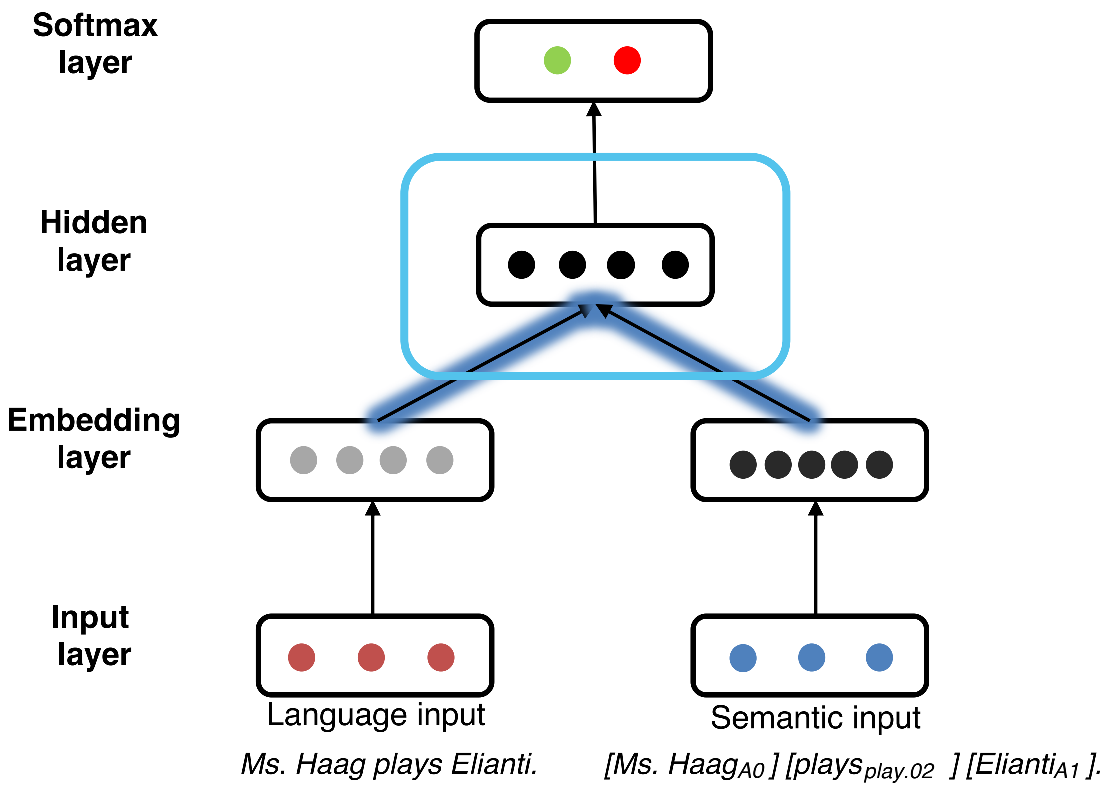
## Embedding Layer

- **Language embedding:** $$\mathcal{L}_{\vec{e}_L} = \sum_{i=1}^{T} \sum_{-c \le j-i \le c, j-i \ne 0} \log q(w_j|w_i)$$ $$q(w_j|w_i) = \frac{\exp(\vec{v}_j'^{T} \vec{v}_i)}{\sum_{k=1}^{|W|} \exp(\vec{v}_k'^{T} \vec{v}_i)}$$
- **Semantic embedding:** $$\mathcal{L}_{\vec{e}_S} = \sum_{p \in \mathcal{P}} \sum_{a \in \mathcal{A}} \log(q(a|p) + q(p|a))$$ $$q(a|p) = \frac{\exp\{z(\vec{p}, \vec{a})\}}{\sum_{a' \in \mathcal{A}} \exp\{z(\vec{p}, \vec{a'})\}}$$ $$z(\vec{p}, \vec{a}) = b - \frac{1}{2}||\vec{p} - \vec{a}||^2$$



## Hidden Layer: Joint Language and Semantic Embedding

- Rectified Linear Units (ReLU): $\vec{h} = \max(0, \mathbf{W}^{e_L h}\vec{e}_L + \mathbf{W}^{e_S h}\vec{e}_S + \vec{b}^h)$

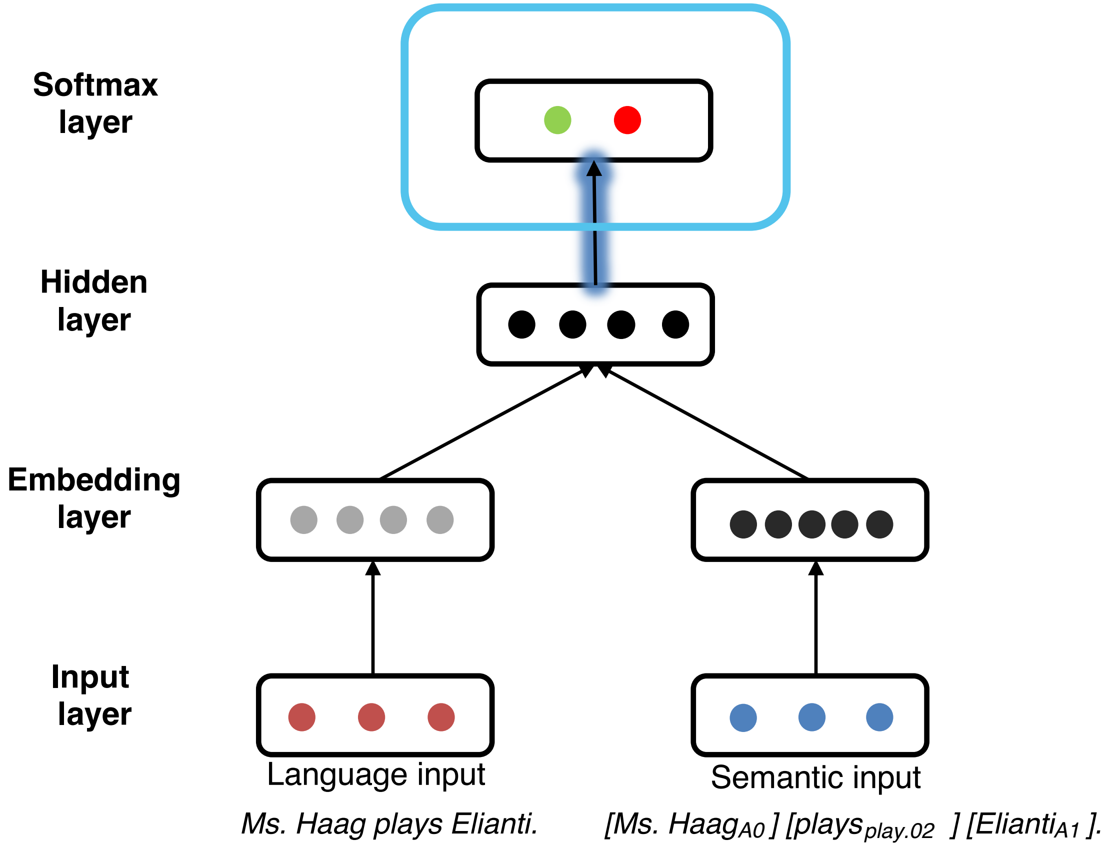


# **Core in Our Approach:** Neural Query Strategy Model

## Softmax Layer

$$q(y=n|\vec{h}) = \frac{e^{\mathbf{W}_n^{e_L s}\vec{e}_L + \mathbf{W}_n^{e_S s}\vec{e}_S + \mathbf{W}_n^{hs}\vec{h} + \vec{b}_n^s}}{\sum_{k=1}^{|K|} e^{\mathbf{W}_k^{e_L s}\vec{e}_L + \mathbf{W}_k^{e_S s}\vec{e}_S + \mathbf{W}_k^{hs}\vec{h} + \vec{b}_k^s}}$$

- **Human-free SRL** label
- **Human-need SRL** label



# Results

| Method | Setting | TEST$_{id}$ P | TEST$_{id}$ R | TEST$_{id}$ F1 | TEST$_{od}$ P | TEST$_{od}$ R | TEST$_{od}$ F1 |
|---|---|---|---|---|---|---|---|
| MATE | Initial | 86.11 | 81.11 | 83.53 | 75.70 | 68.38 | 71.86 |
| | ACTIVESRL$_{nf}$ | 87.07 | 82.45 | 84.70 | 76.25 | 70.47 | 73.25 |
| | ACTIVESRL | **87.69** | **83.42** | **85.50** | **76.79** | **71.27** | **73.93** |
| | Upper Bound | 89.59 | 86.07 | 87.79 | 79.46 | 74.21 | 76.74 |
| CLEAR | Initial | 82.07 | 70.57 | 75.89 | 72.77 | 62.14 | 67.09 |
| | ACTIVESRL$_{nf}$ | 83.12 | 72.97 | 77.72 | 73.02 | 62.57 | 67.44 |
| | ACTIVESRL | **83.65** | **73.74** | **78.38** | **74.37** | **66.90** | **70.48** |
| | Upper Bound | 84.74 | 74.47 | 79.27 | 75.44 | 67.20 | 71.08 |
| K-SRL | Initial | 89.54 | 80.50 | 84.78 | 81.39 | 69.34 | 74.88 |
| | ACTIVESRL$_{nf}$ | 90.37 | 82.90 | 86.48 | 82.15 | 71.27 | 76.33 |
| | ACTIVESRL | **91.05** | **84.44** | **87.62** | **82.67** | **72.74** | **77.39** |
| | Upper Bound | 91.21 | 87.42 | 89.28 | 82.09 | 77.84 | 79.91 |

**Effective for popular SRL models**

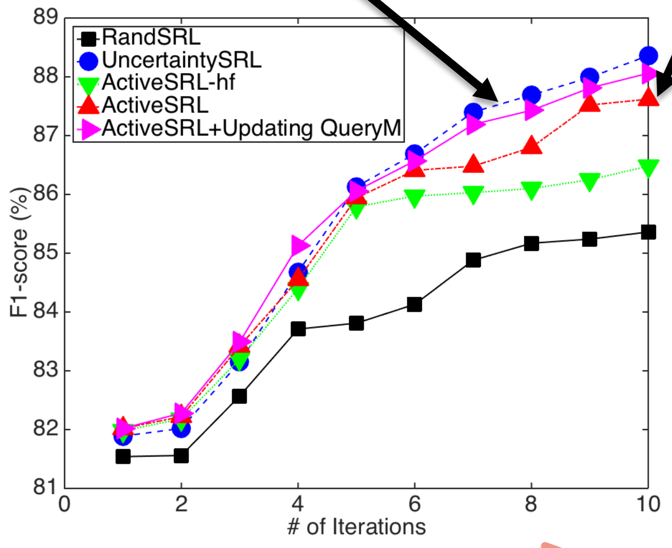
# Results

| Method | Setting | TEST$_{id}$ P | TEST$_{id}$ R | TEST$_{id}$ F1 | TEST$_{od}$ P | TEST$_{od}$ R | TEST$_{od}$ F1 |
|---|---|---|---|---|---|---|---|
| MATE | Initial | 86.11 | 81.11 | 83.53 | 75.70 | 68.38 | 71.86 |
| | ACTIVESRL$_{hf}$ | 87.07 | 82.45 | 84.70 | 76.25 | 70.47 | 73.25 |
| | ACTIVESRL | **87.69** | **83.42** | **85.50** | **76.79** | **71.27** | **73.93** |
| | Upper Bound | 89.59 | 86.07 | 87.79 | 79.46 | 74.21 | 76.74 |
| CLEAR | Initial | 82.07 | 70.57 | 75.89 | 72.77 | 62.14 | 67.09 |
| | ACTIVESRL$_{hf}$ | 83.12 | 72.97 | 77.72 | 73.02 | 62.57 | 67.44 |
| | ACTIVESRL | **83.65** | **73.74** | **78.38** | **74.37** | **66.90** | **70.48** |
| | Upper Bound | 84.74 | 74.47 | 79.27 | 75.44 | 67.20 | 71.08 |
| K-SRL | Initial | 89.54 | 80.50 | 84.78 | 81.39 | 69.34 | 74.88 |
| | ACTIVESRL$_{hf}$ | 90.37 | 82.90 | 86.48 | 82.15 | 71.27 | 76.33 |
| | ACTIVESRL | **91.05** | **84.44** | **87.62** | **82.67** | **72.74** | **77.39** |
| | Upper Bound | 91.21 | 87.42 | 89.28 | 82.09 | 77.84 | 79.91 |

**Effective for popular SRL models**

Active learning for **white-box SRL**

Active learning for **black-box SRL**




**Comparable to active learning for white-box SRL (model details are known)**

# Take Away

- Active learning framework for black-box SRL models
- Neural query strategy model to learn the strategy for selecting the data instances

- Active learning framework for black-box SRL models
- Neural query strategy model to learn the strategy for selecting the data instances

More importantly, if you have no knowledge about the model, or you are too lazy to design a query strategy for active learning, just try our approach☺