Guiding Neural Entity Alignment with Compatibility

Bing Liu[1], Harrisen Scells[1], Wen Hua[1], Guido Zuccon[1], Genghong Zhao[2], Xia Zhang[2]

The University of Queensland[1], Neusoft[2]

bing.liu@uq.edu.au

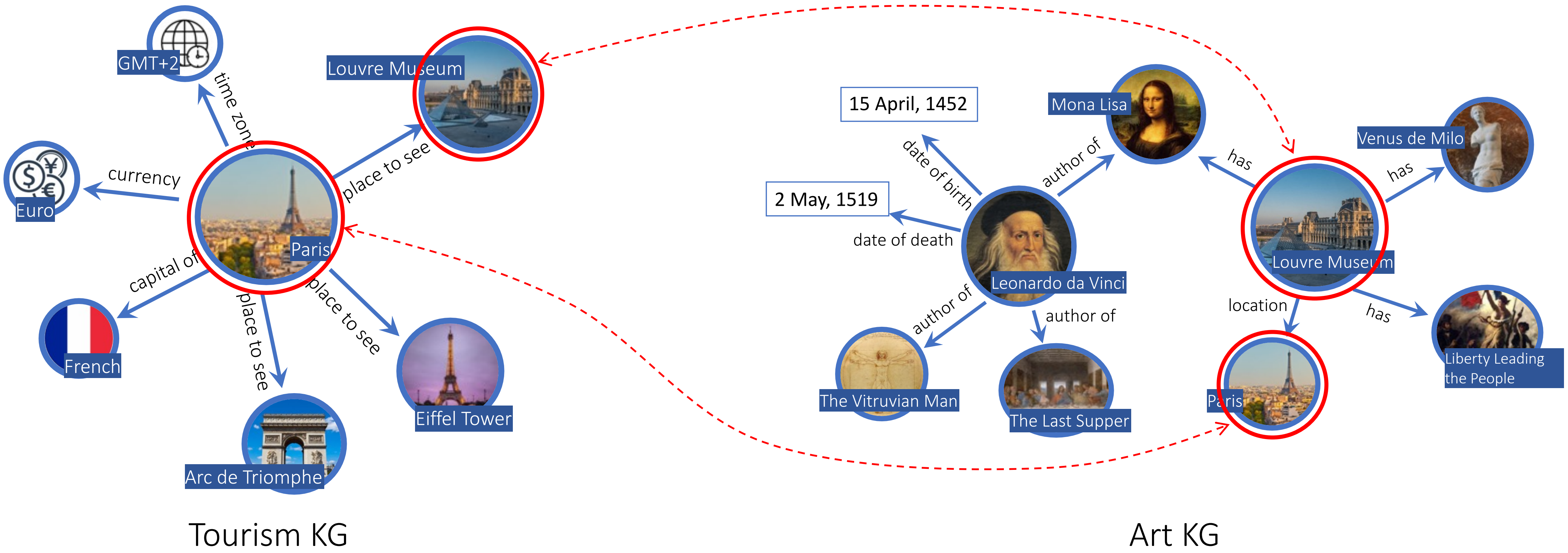
https://uqbingliu.github.io/

@BingLiu1011

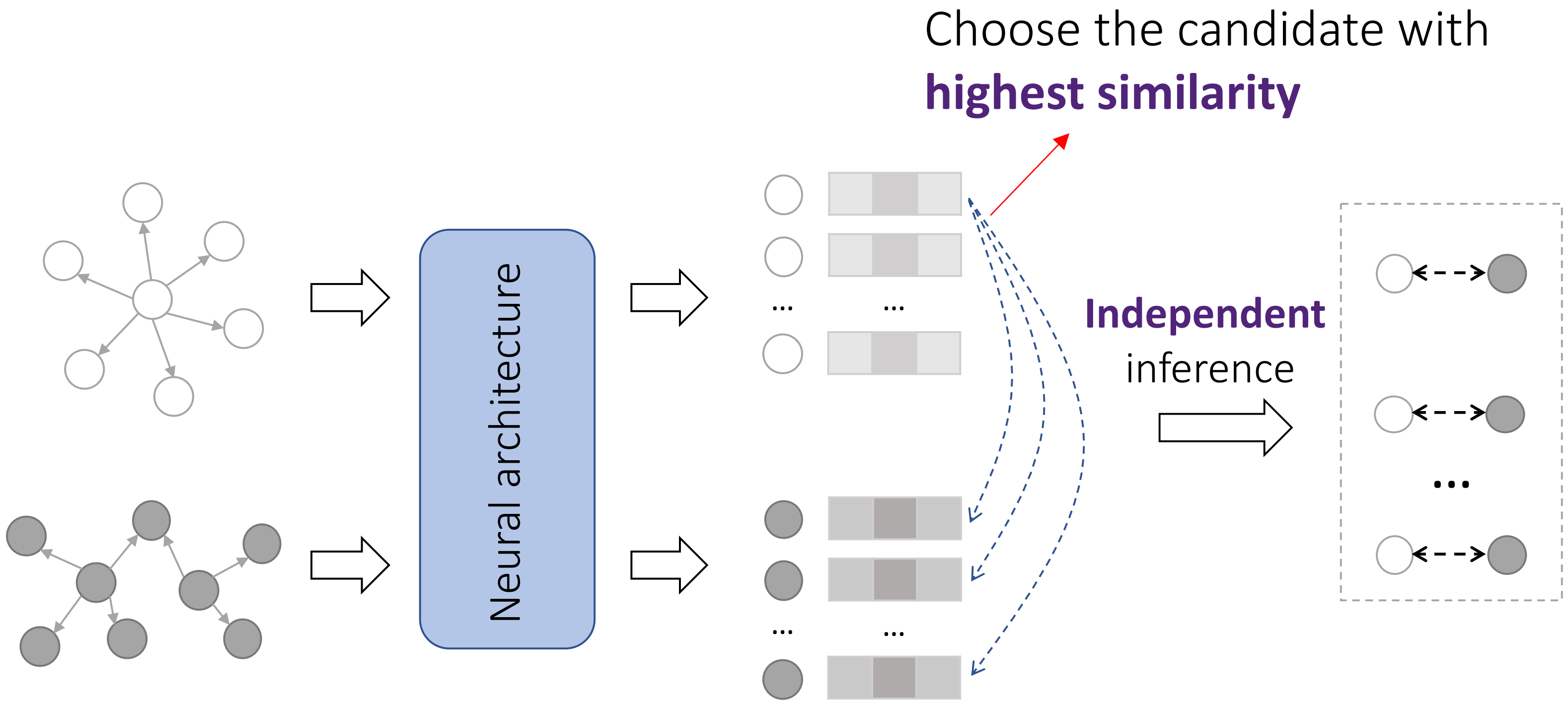
Entity Alignment (EA)

Establish **mappings** between **equivalent entities** in different Knowledge Graphs (KGs).

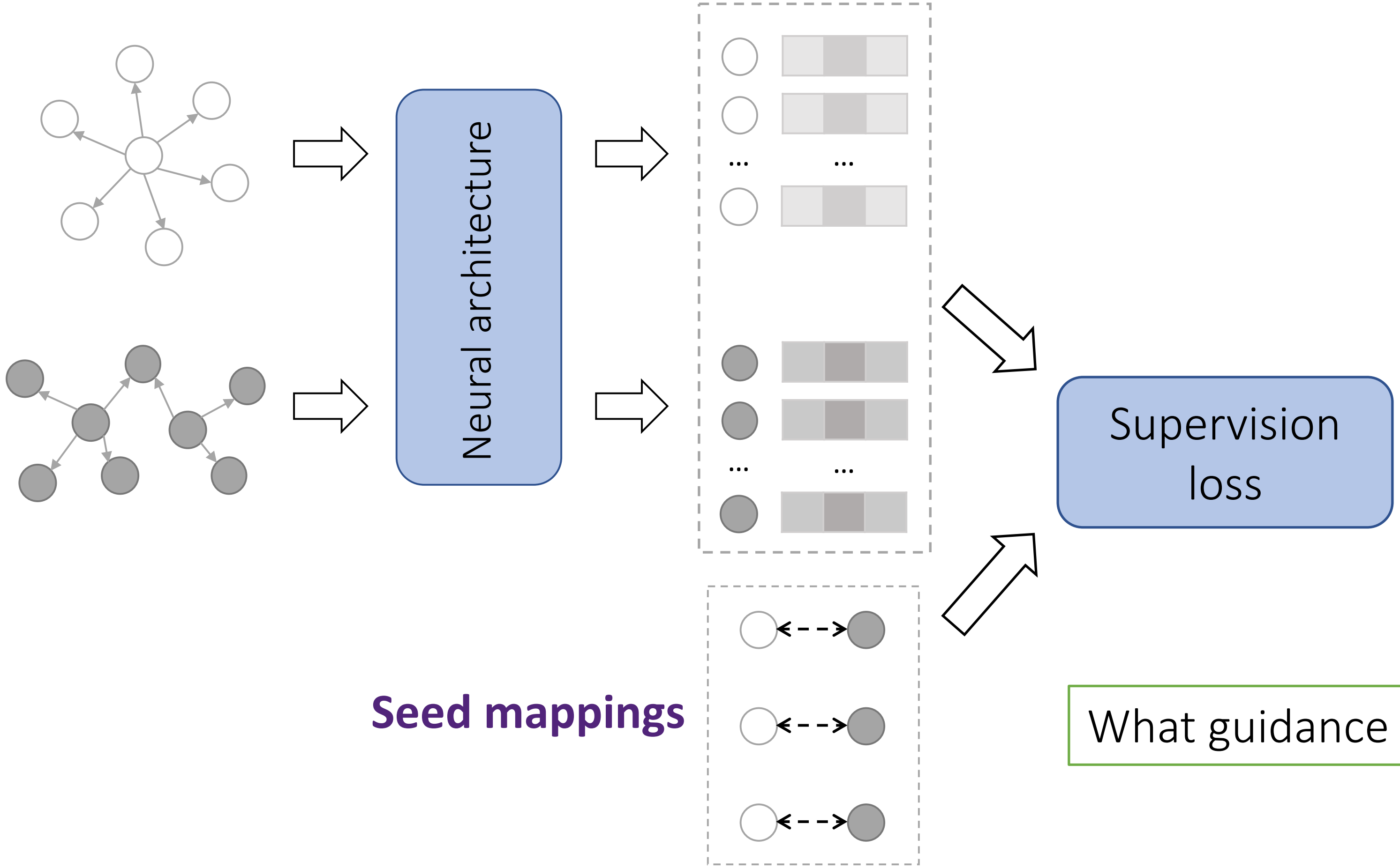
- Challenge: **Heterogeneity** of different KGs



Neural Entity Alignment

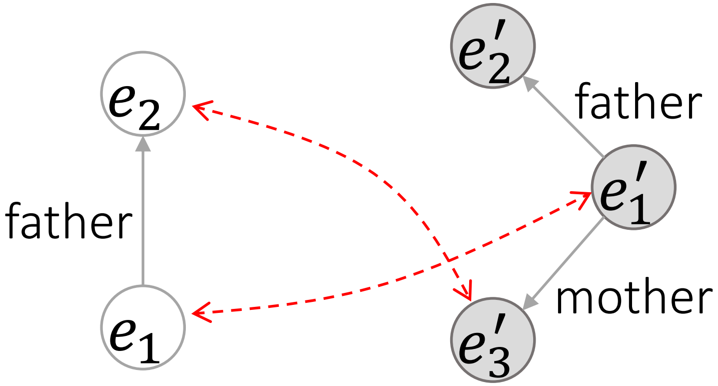


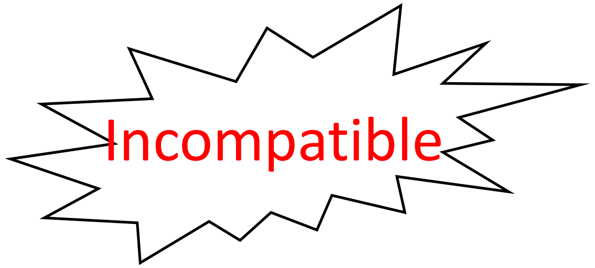
Neural Entity Alignment - Training



What guidance else can we use?

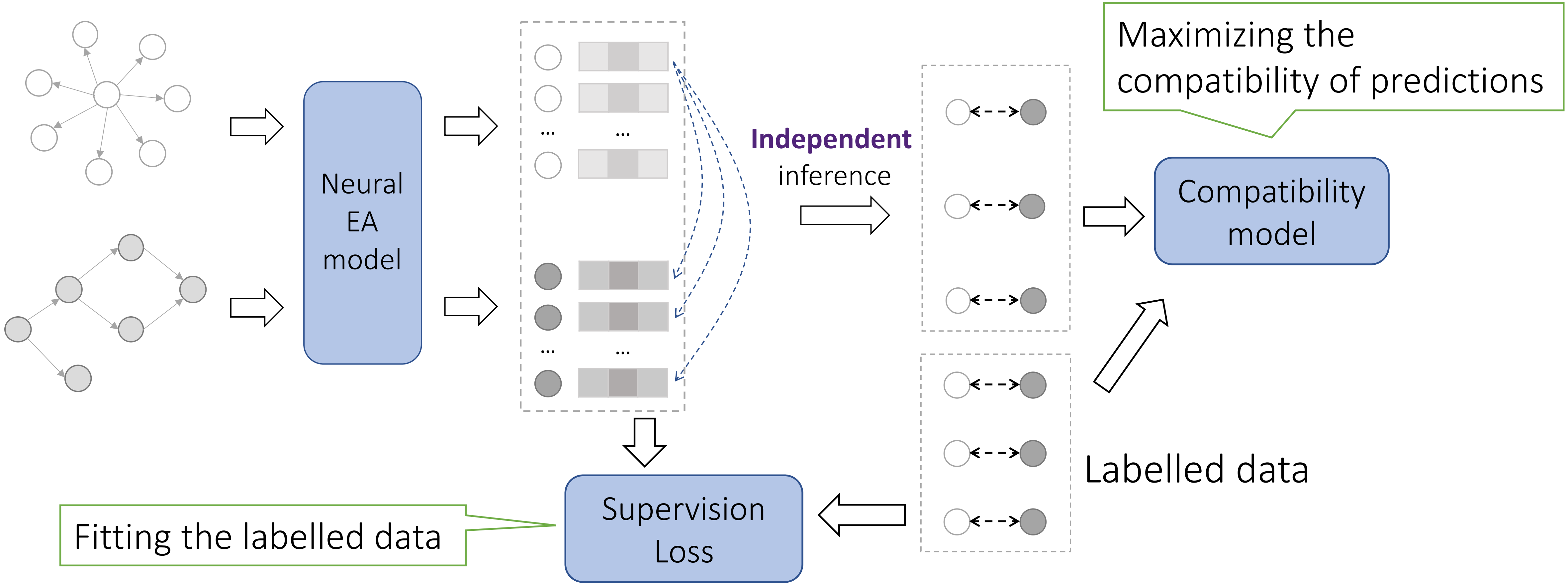
Compatibility of Mappings





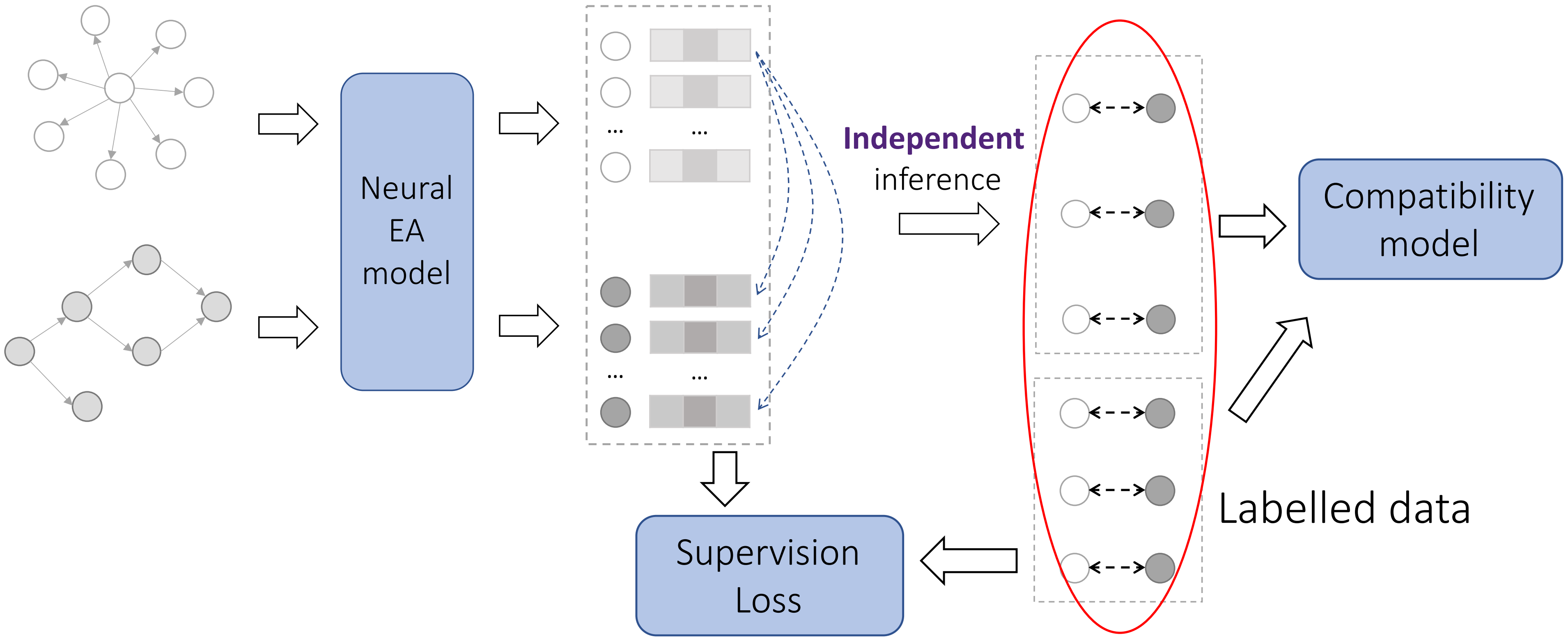
Another Objective: Making Compatible Predictions

Learning objectives:



Challenges

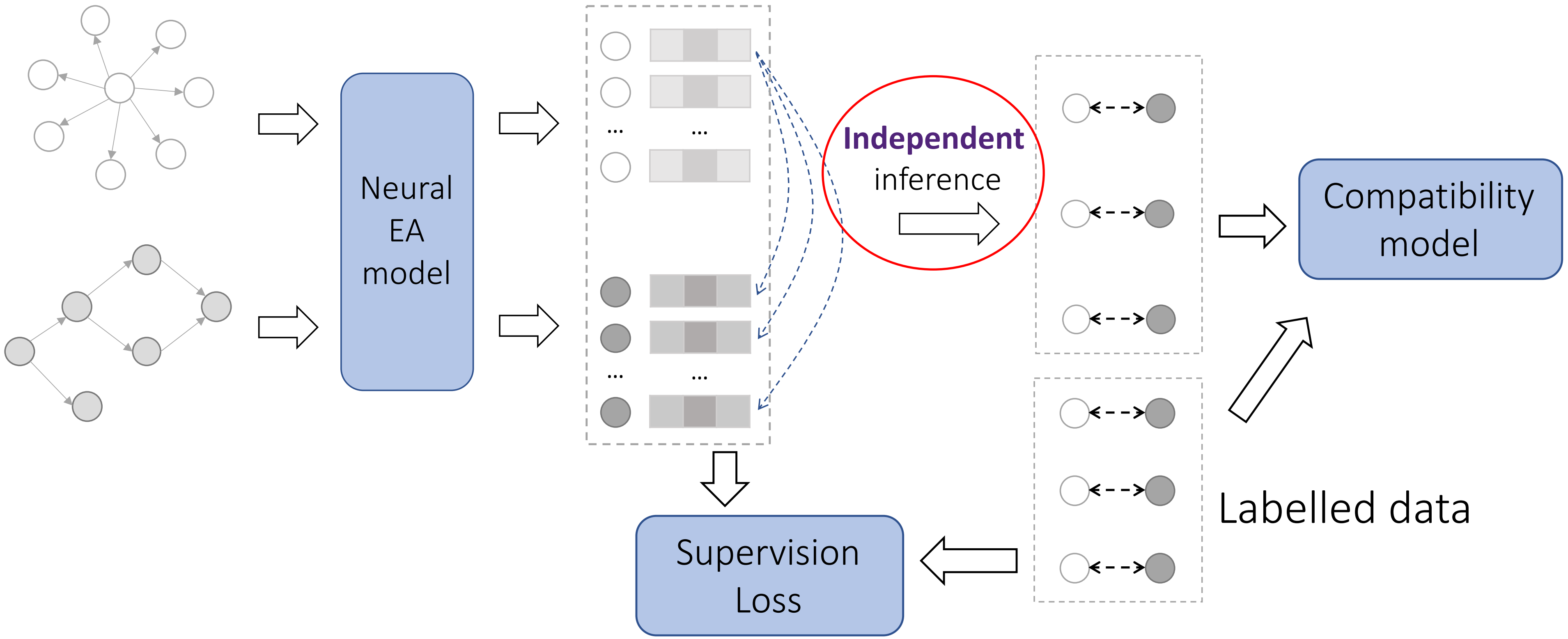
Challenge 1: How to measure the overall compatibility of a large number of mappings?



Challenges

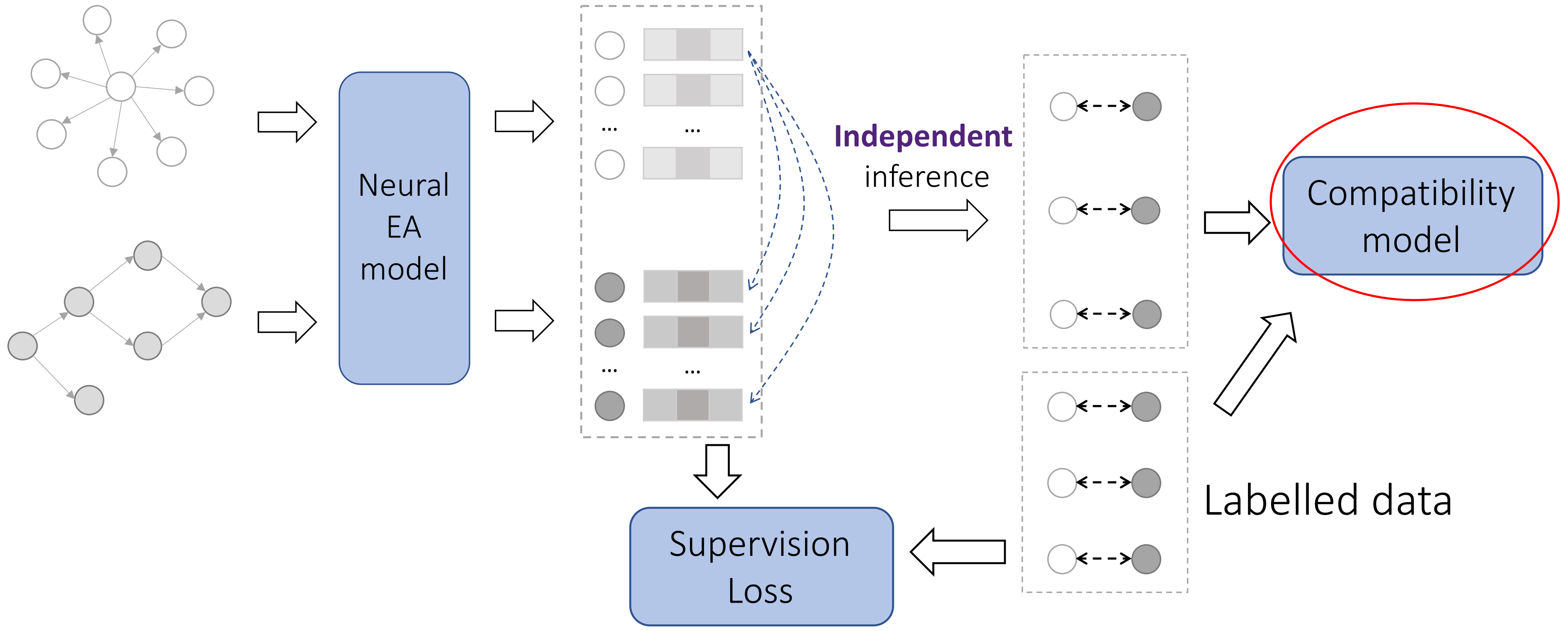
Challenge 2: How to exploit the compatibility to train the EA models ?

- The inference processing is not derivable.



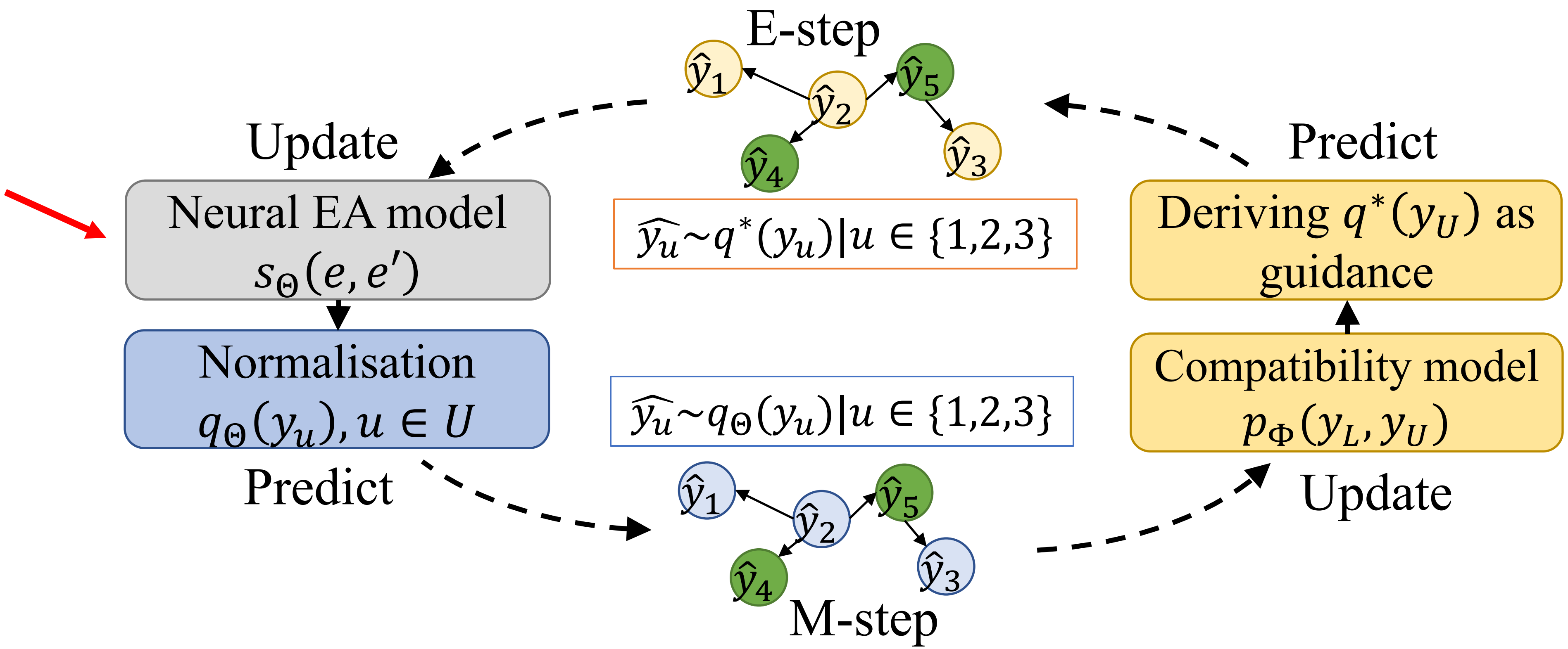
Challenges

Challenge 3: How to optimize the compatibility model?



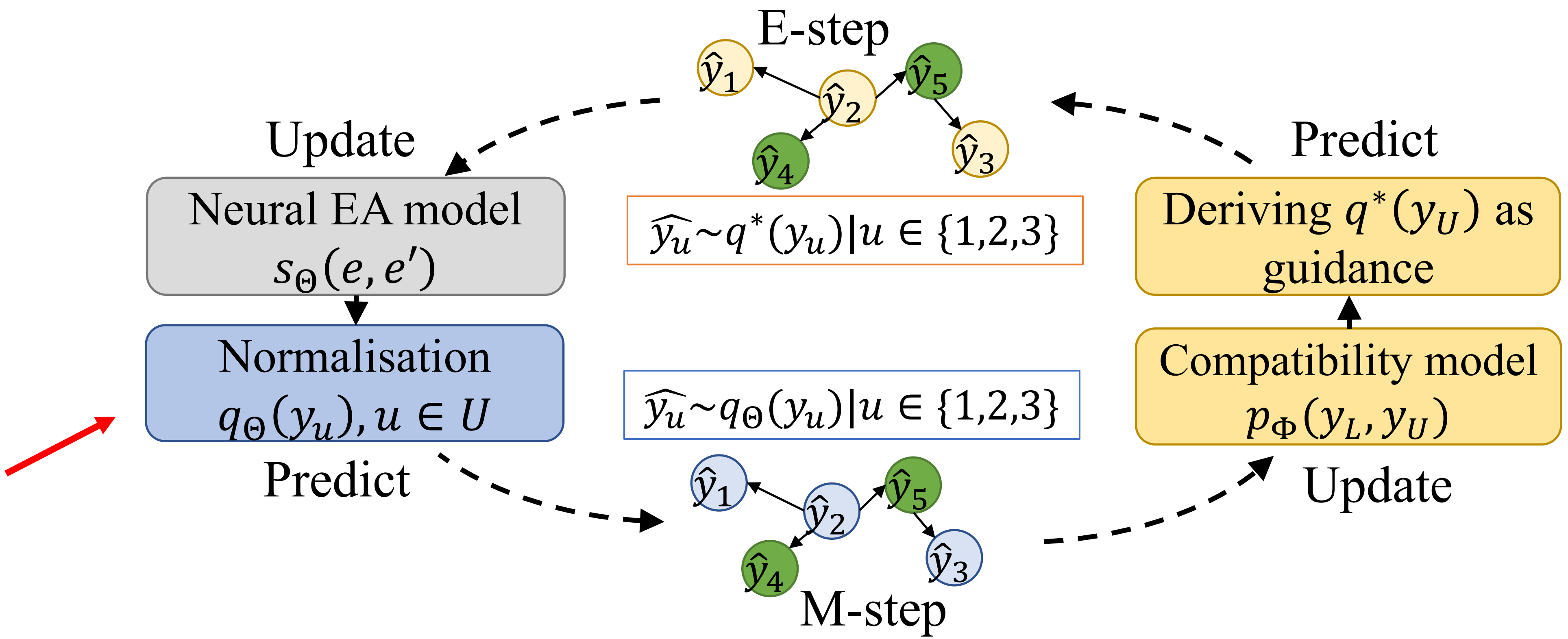
The EMEA Framework

Overview

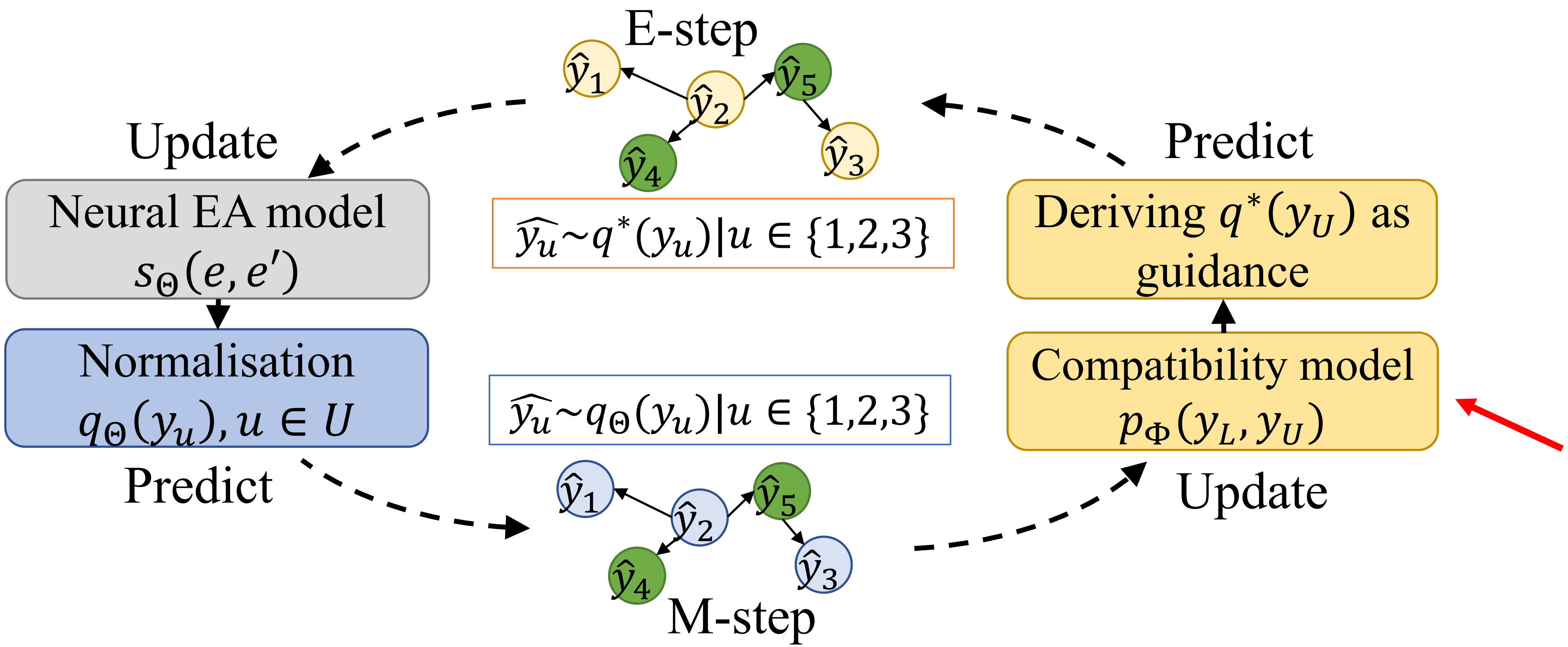


E-step
$\hat{y}_1$ $\hat{y}_2$ $\hat{y}_5$ $\hat{y}_4$ $\hat{y}_3$
Update
Predict
Neural EA model
$s_\Theta(e, e')$
$\widehat{y_u} \sim q^*(y_u) | u \in \{1,2,3\}$
Deriving $q^*(y_U)$ as guidance
Normalisation
$q_\Theta(y_u), u \in U$
$\widehat{y_u} \sim q_\Theta(y_u) | u \in \{1,2,3\}$
Compatibility model
$p_\Phi(y_L, y_U)$
Predict
Update
$\hat{y}_1$ $\hat{y}_2$ $\hat{y}_5$ $\hat{y}_4$ $\hat{y}_3$
M-step

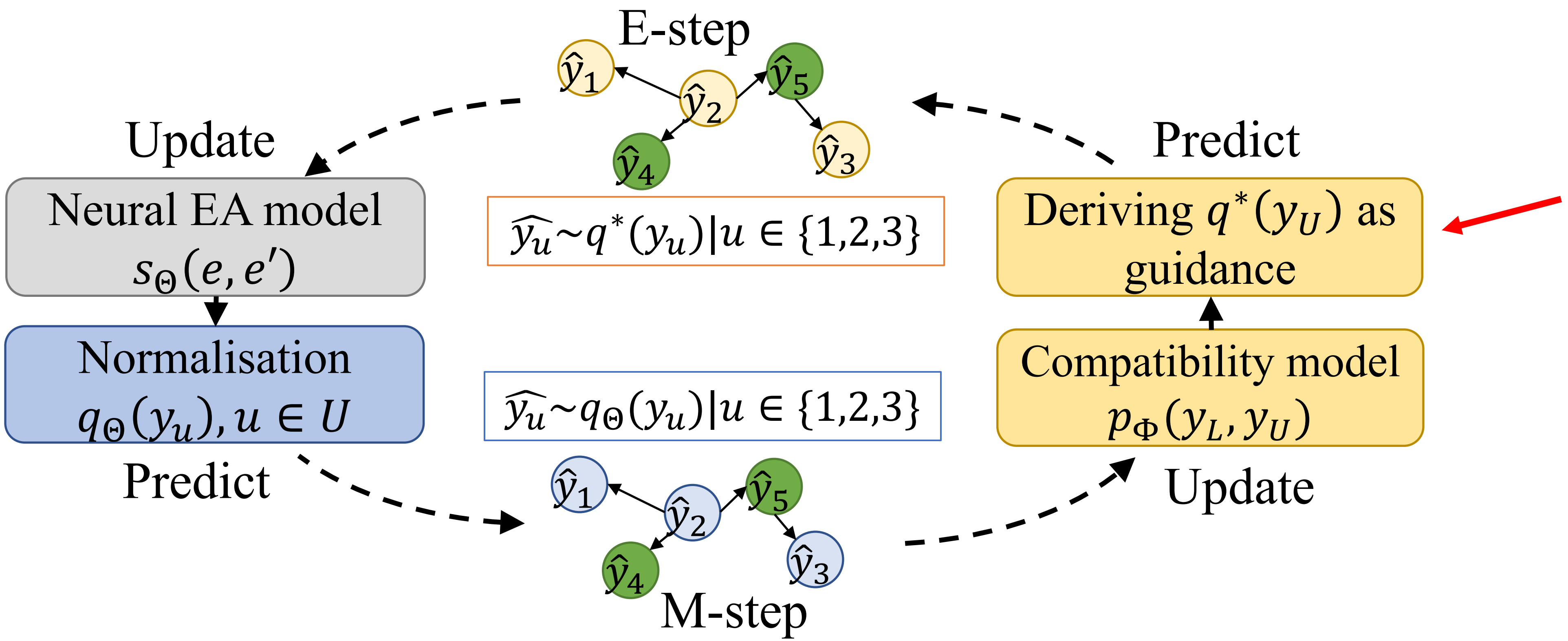
Overview


E-step
$\hat{y}_1$
$\hat{y}_2$
$\hat{y}_3$
$\hat{y}_4$
$\hat{y}_5$
Update
Predict
Neural EA model
$s_\Theta(e, e')$
$\widehat{y_u} \sim q^*(y_u) | u \in \{1,2,3\}$
Deriving $q^*(y_U)$ as guidance
Normalisation
$q_\Theta(y_u), u \in U$
$\widehat{y_u} \sim q_\Theta(y_u) | u \in \{1,2,3\}$
Compatibility model
$p_\Phi(y_L, y_U)$
Predict
Update
$\hat{y}_1$
$\hat{y}_2$
$\hat{y}_3$
$\hat{y}_4$
$\hat{y}_5$
M-step

Overview

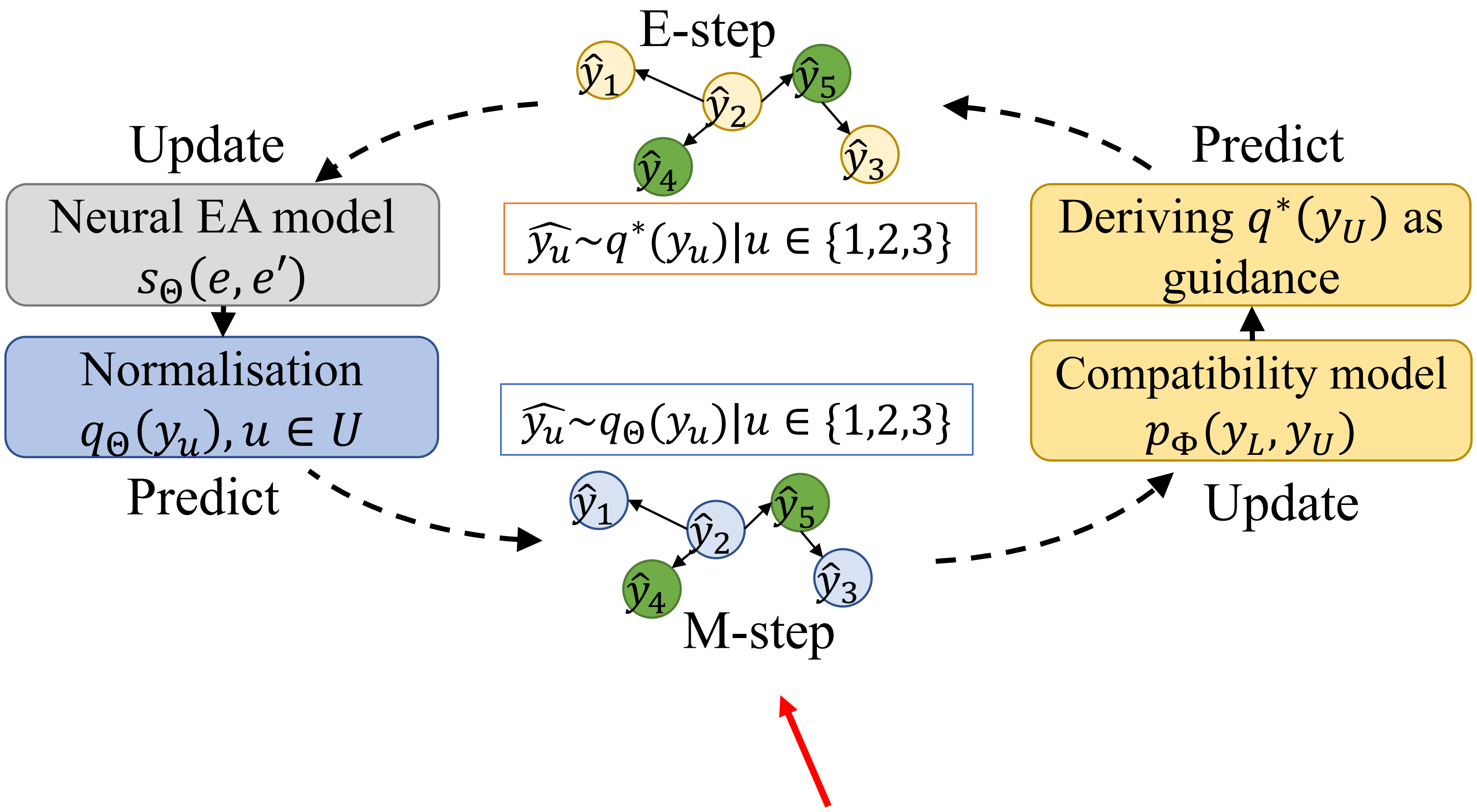

E-step
$\hat{y}_1$ $\hat{y}_2$ $\hat{y}_5$ $\hat{y}_4$ $\hat{y}_3$
Update
Predict
Neural EA model
$s_\Theta(e, e')$
$\widehat{y_u} \sim q^*(y_u) | u \in \{1,2,3\}$
Deriving $q^*(y_U)$ as guidance
Normalisation
$q_\Theta(y_u), u \in U$
$\widehat{y_u} \sim q_\Theta(y_u) | u \in \{1,2,3\}$
Compatibility model
$p_\Phi(y_L, y_U)$
Predict
Update
$\hat{y}_1$ $\hat{y}_2$ $\hat{y}_5$ $\hat{y}_4$ $\hat{y}_3$
M-step

Overview

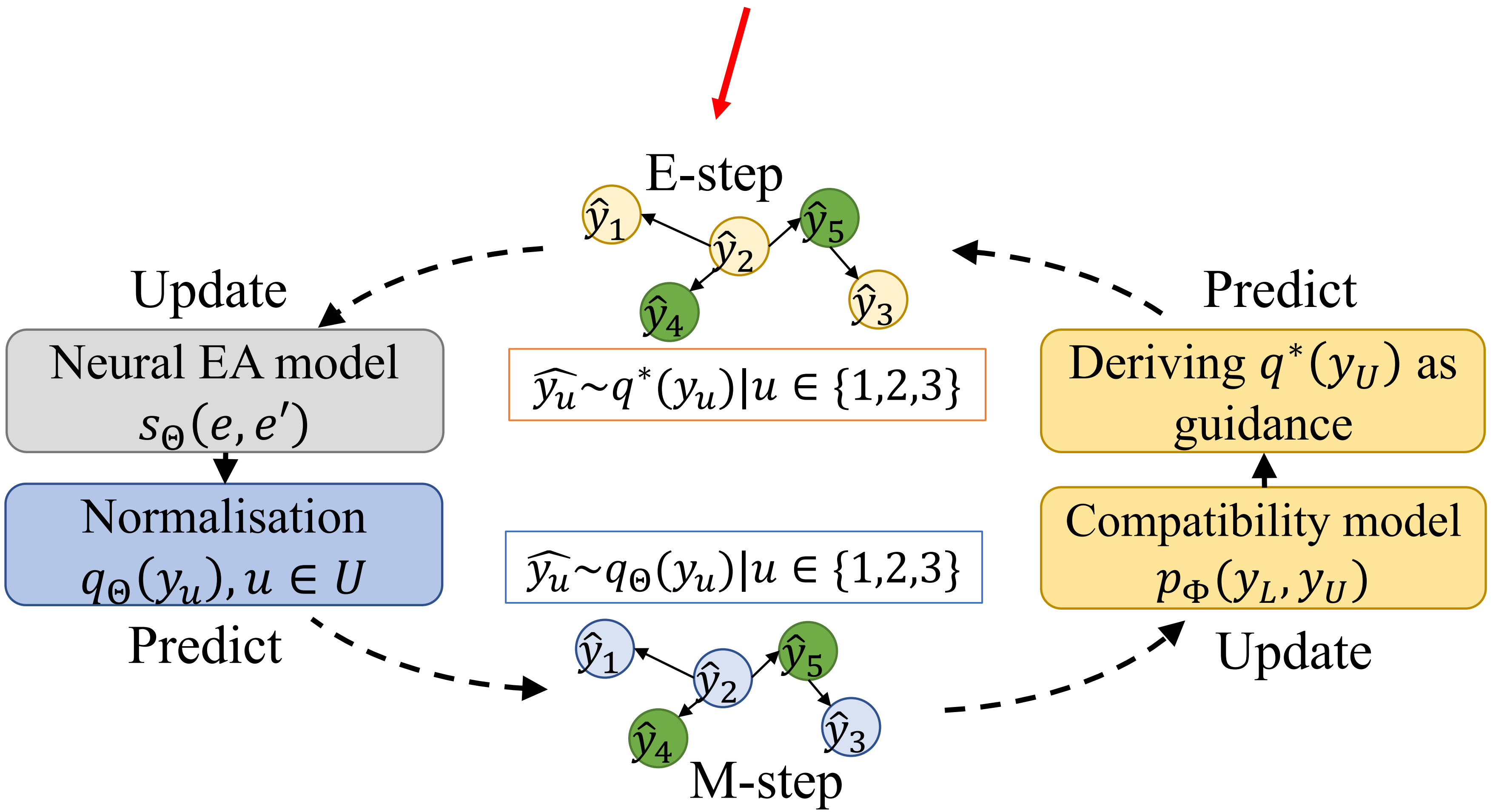


E-step
$\hat{y}_1$
$\hat{y}_2$
$\hat{y}_5$
$\hat{y}_4$
$\hat{y}_3$
Update
Predict
Neural EA model $s_\Theta(e, e')$
$\widehat{y_u} \sim q^*(y_u) | u \in \{1,2,3\}$
Deriving $q^*(y_U)$ as guidance
Normalisation $q_\Theta(y_u), u \in U$
$\widehat{y_u} \sim q_\Theta(y_u) | u \in \{1,2,3\}$
Compatibility model $p_\Phi(y_L, y_U)$
Predict
Update
$\hat{y}_1$
$\hat{y}_2$
$\hat{y}_5$
$\hat{y}_4$
$\hat{y}_3$
M-step

Overview


E-step
$\hat{y}_1$ $\hat{y}_2$ $\hat{y}_3$ $\hat{y}_4$ $\hat{y}_5$
Update
Neural EA model
$s_\Theta(e, e')$
$\widehat{y_u} \sim q^*(y_u) | u \in \{1,2,3\}$
Predict
Deriving $q^*(y_U)$ as guidance
Normalisation
$q_\Theta(y_u), u \in U$
$\widehat{y_u} \sim q_\Theta(y_u) | u \in \{1,2,3\}$
Compatibility model
$p_\Phi(y_L, y_U)$
Predict
$\hat{y}_1$ $\hat{y}_2$ $\hat{y}_3$ $\hat{y}_4$ $\hat{y}_5$
Update
M-step

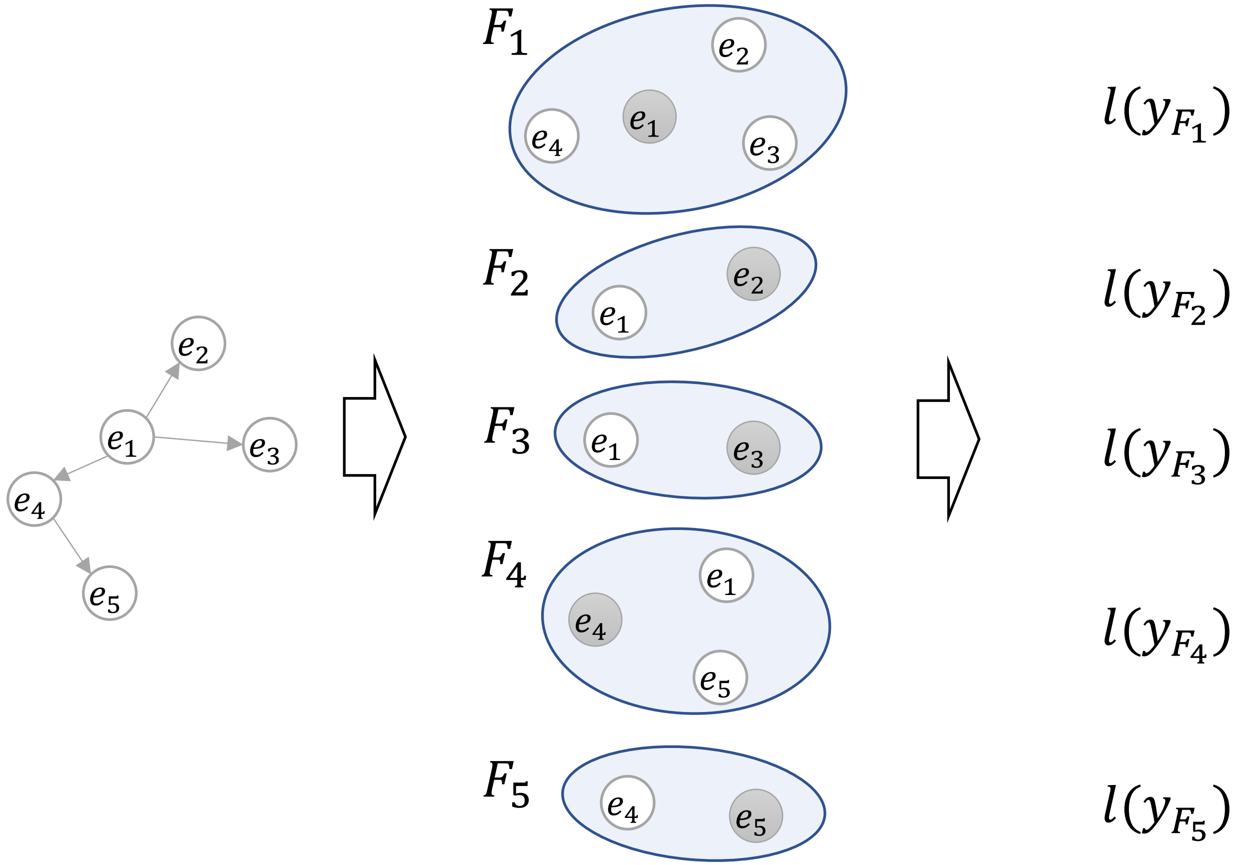
Overview


E-step
$\hat{y}_1$
$\hat{y}_2$
$\hat{y}_5$
$\hat{y}_4$
$\hat{y}_3$
Update
Predict
Neural EA model
$s_\Theta(e, e')$
$\widehat{y_u} \sim q^*(y_u) | u \in \{1,2,3\}$
Deriving $q^*(y_U)$ as guidance
Normalisation
$q_\Theta(y_u), u \in U$
$\widehat{y_u} \sim q_\Theta(y_u) | u \in \{1,2,3\}$
Compatibility model
$p_\Phi(y_L, y_U)$
Predict
Update
$\hat{y}_1$
$\hat{y}_2$
$\hat{y}_5$
$\hat{y}_4$
$\hat{y}_3$
M-step

Compatibility Model

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

Model the **overall compatibility** with a **Probabilistic Graphical Model**



Source KG **Factor subsets** **Local compatibilities**

Compatibility Model: Local Compatibility

Reuse **PARIS[1] rule** to check compatibility:

- Estimate the equivalence of two entities according to other mappings between their neighbours.

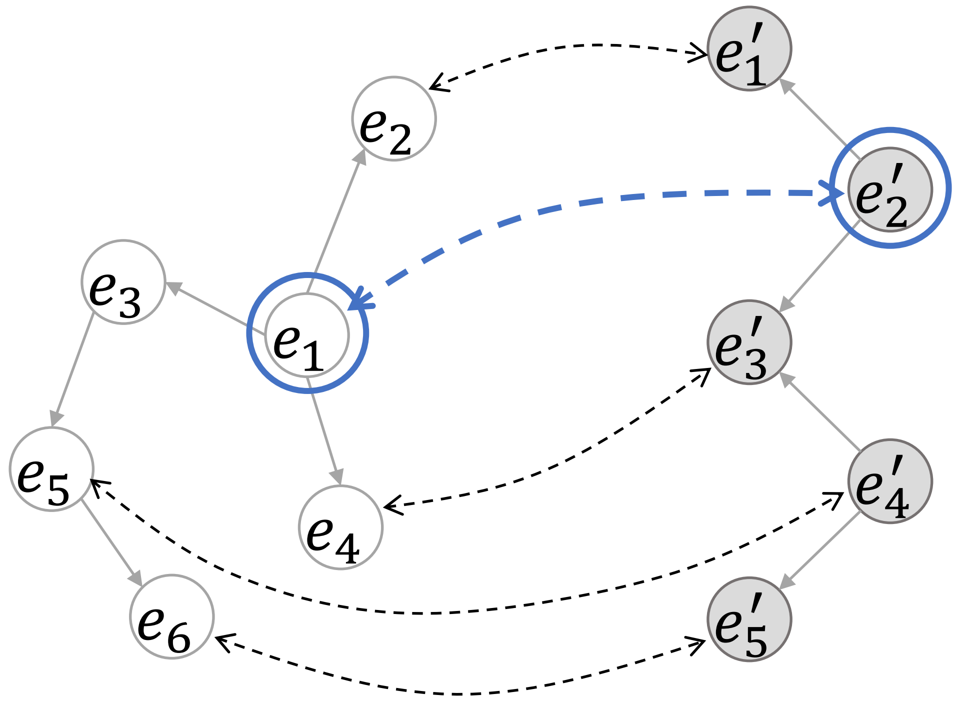
[1] Suchanek et al. "PARIS: Probabilistic Alignment of Relations, Instances, and Schema." Proceedings of the VLDB Endowment 5.3 (2011).

Compatibility Model: Local Compatibility

The University of Queensland Australia

Reuse **PARIS rule** to check compatibility:

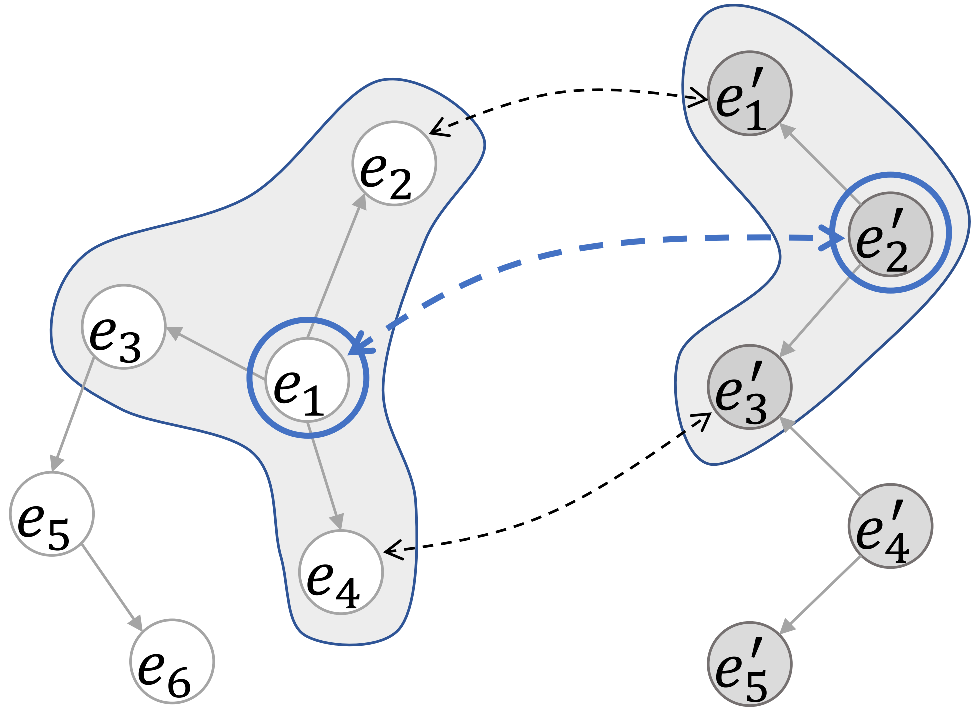
- Estimate the equivalence of two entities according to other mappings between their neighbours.



Compatibility Model: Local Compatibility

Reuse **PARIS rule** to check compatibility:

- Estimate the equivalence of two entities according to other mappings between their neighbours.

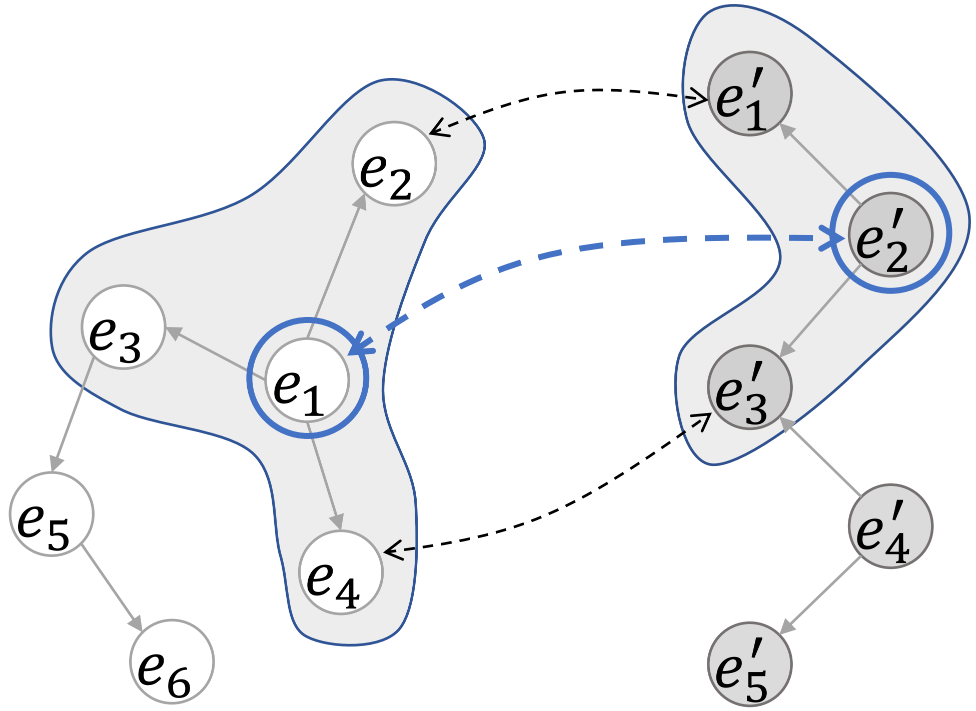


Two supporting mappings

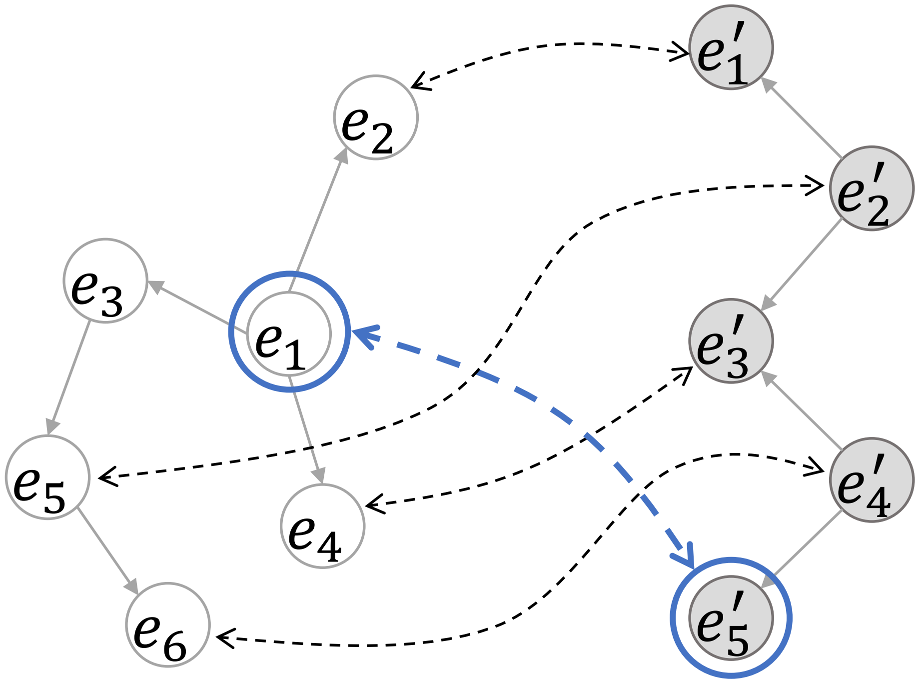
Compatibility Model: Local Compatibility

Reuse **PARIS rule** to check compatibility:

- Estimate the equivalence of two entities according to other mappings between their neighbours.



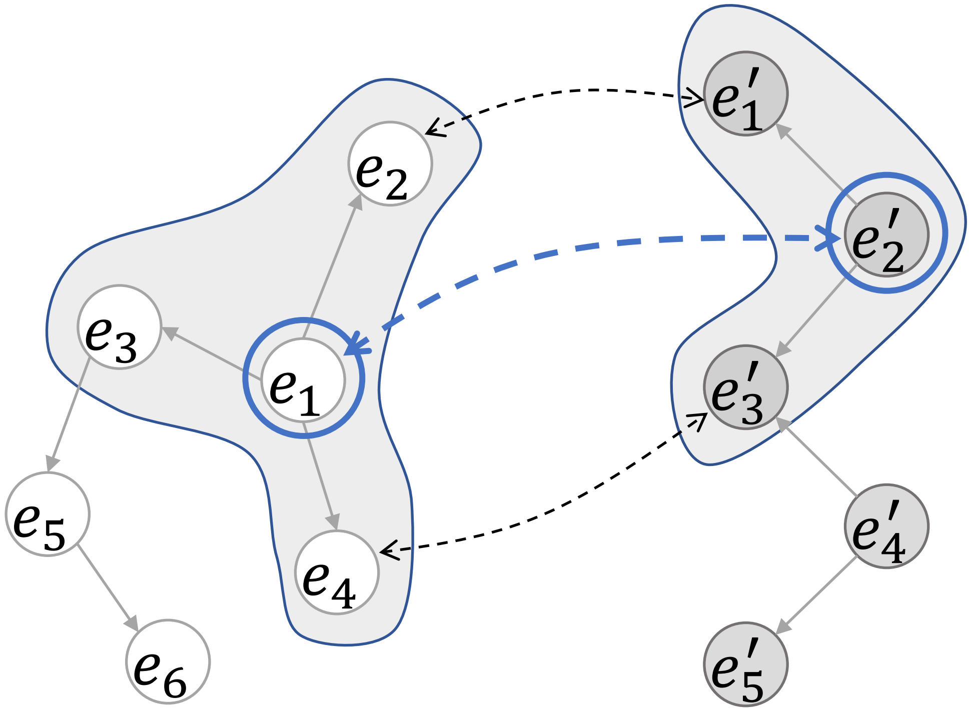
Two supporting mappings



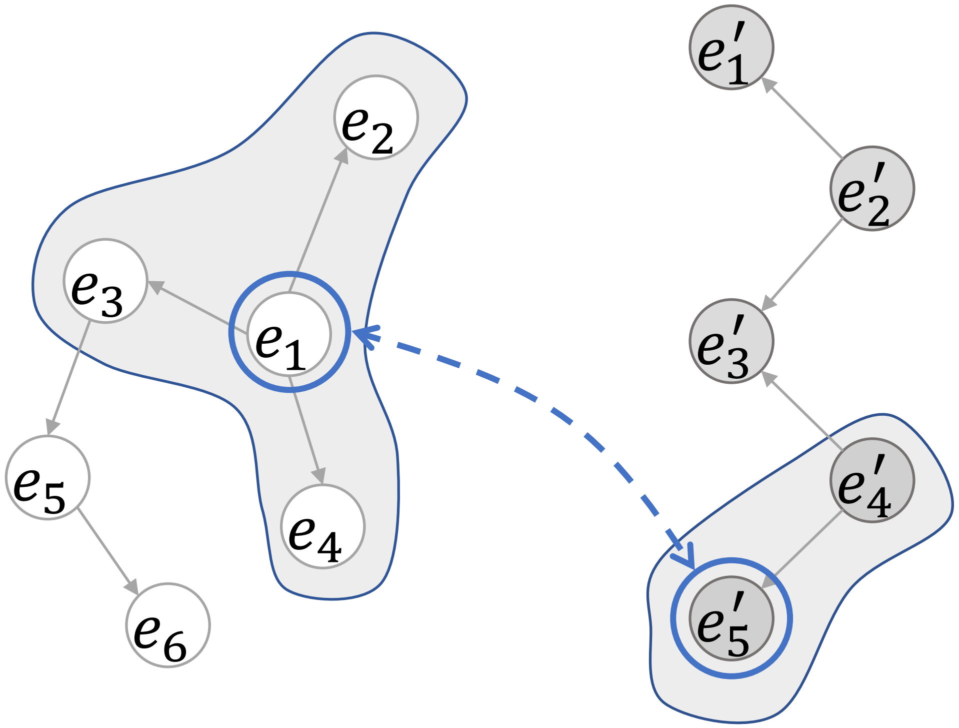
Compatibility Model: Local Compatibility

Reuse **PARIS rule** to check compatibility:

- Estimate the equivalence of two entities according to other mappings between their neighbours.



Two supporting mappings

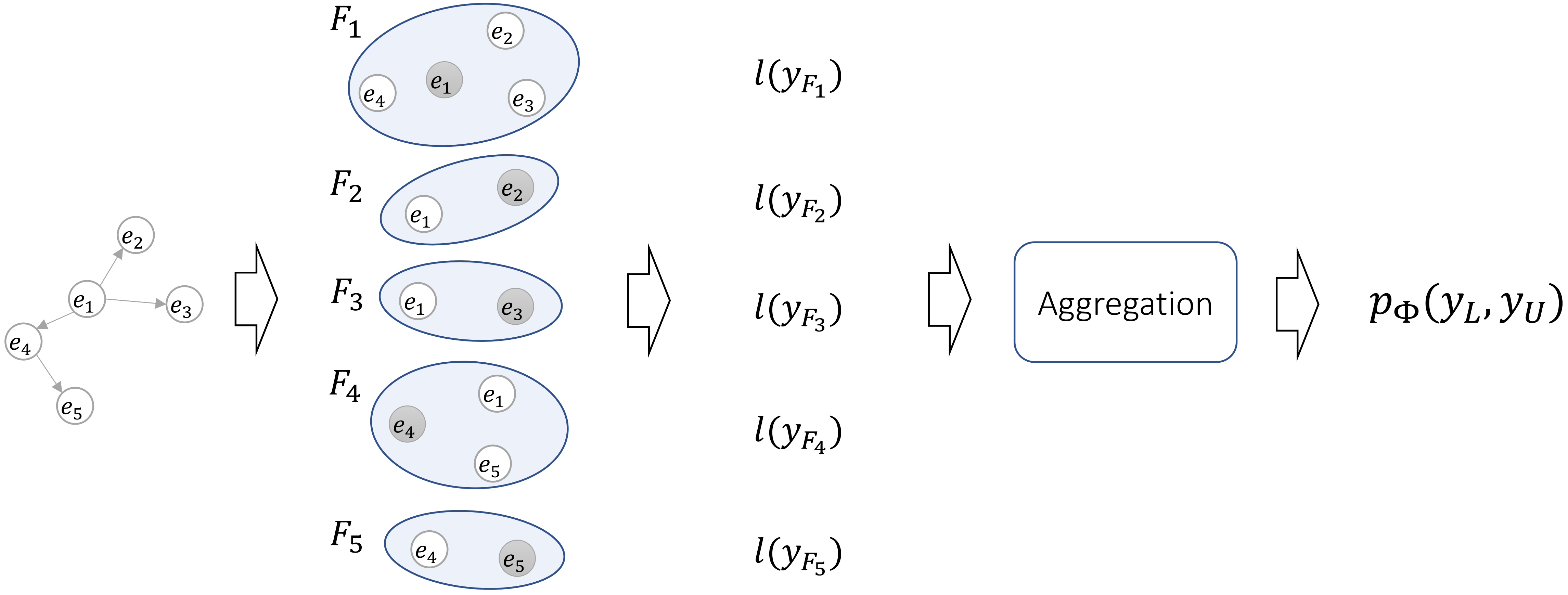


No supporting mapping

Compatibility Model

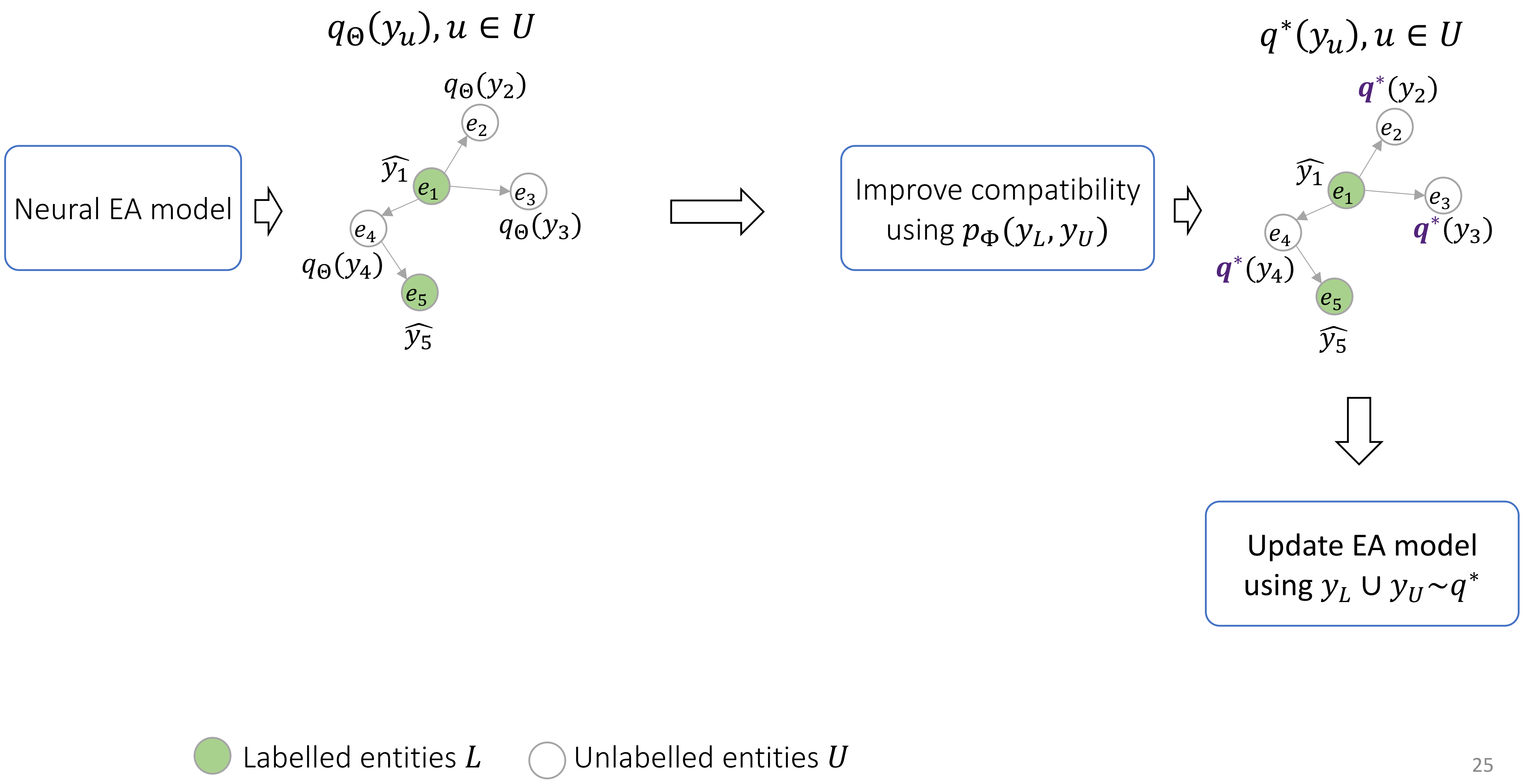
The University of Queensland
Australia

Model the **overall compatibility** with a **Probabilistic Graphical Model**

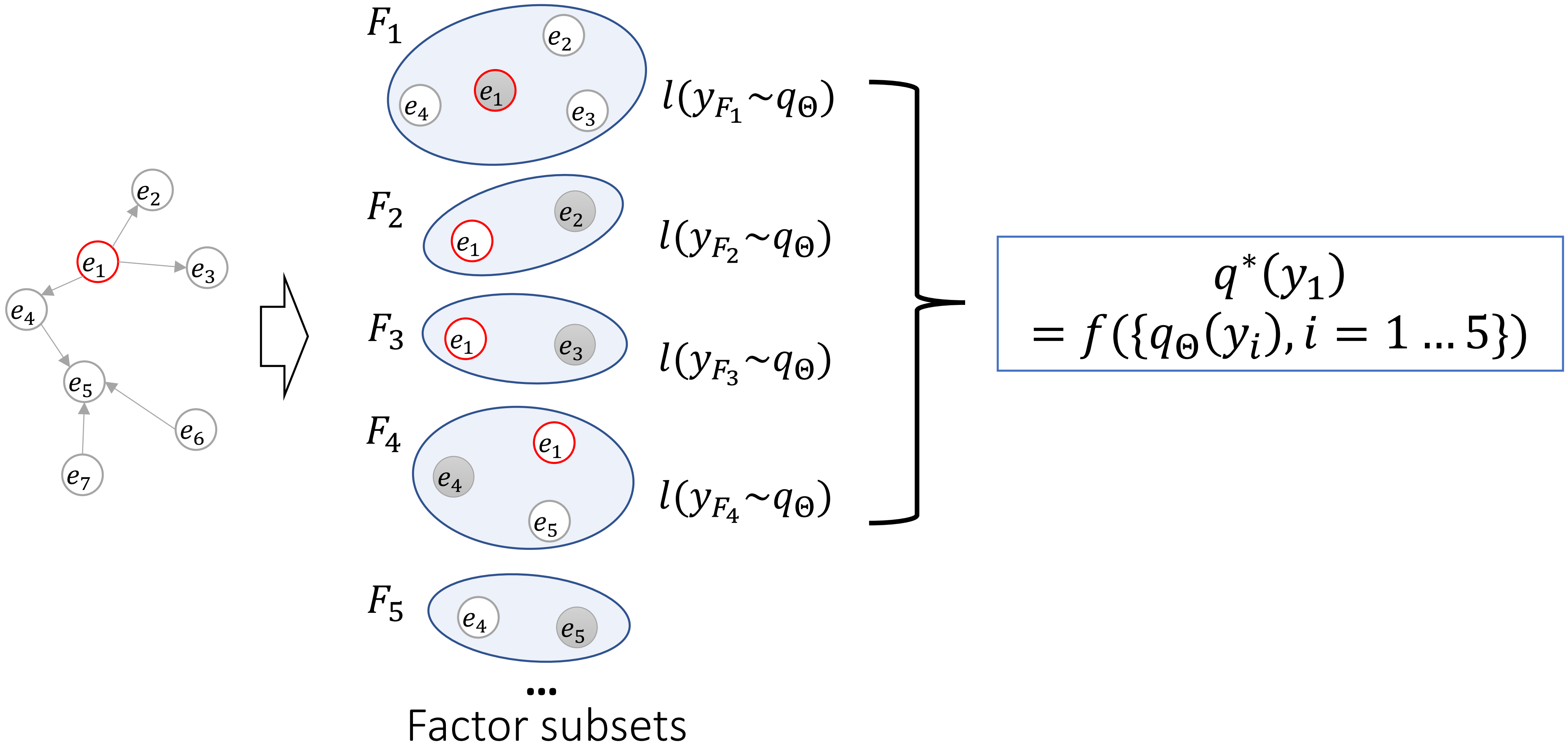


Source KG | **Factor subsets** | **Local compatibilities** | **Joint probability**

Improve Neural Predictions



Improve Neural Predictions

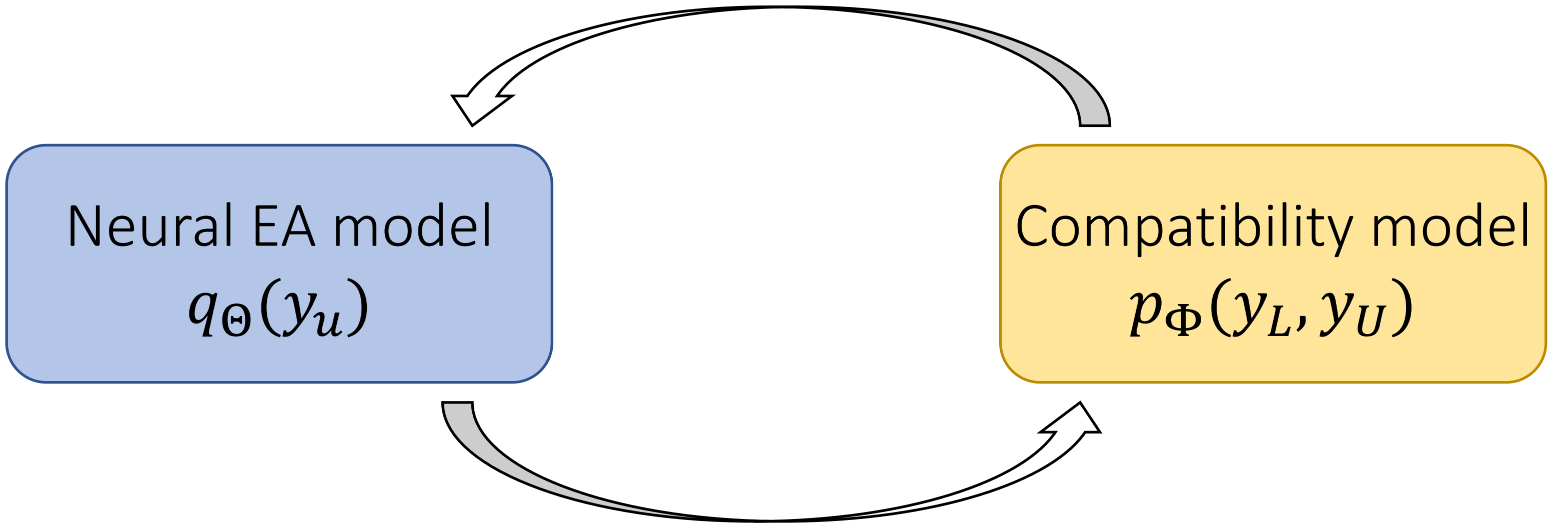


$q^*(y_e)$ involves the neural distributions of entities depending on e.

Optimization with Variational EM

Expectation step:

- Derive $q^*(y_u), u \in U$
- Update Θ using $\{y_u \sim q^*(y_u), u \in U\}$ and y_L

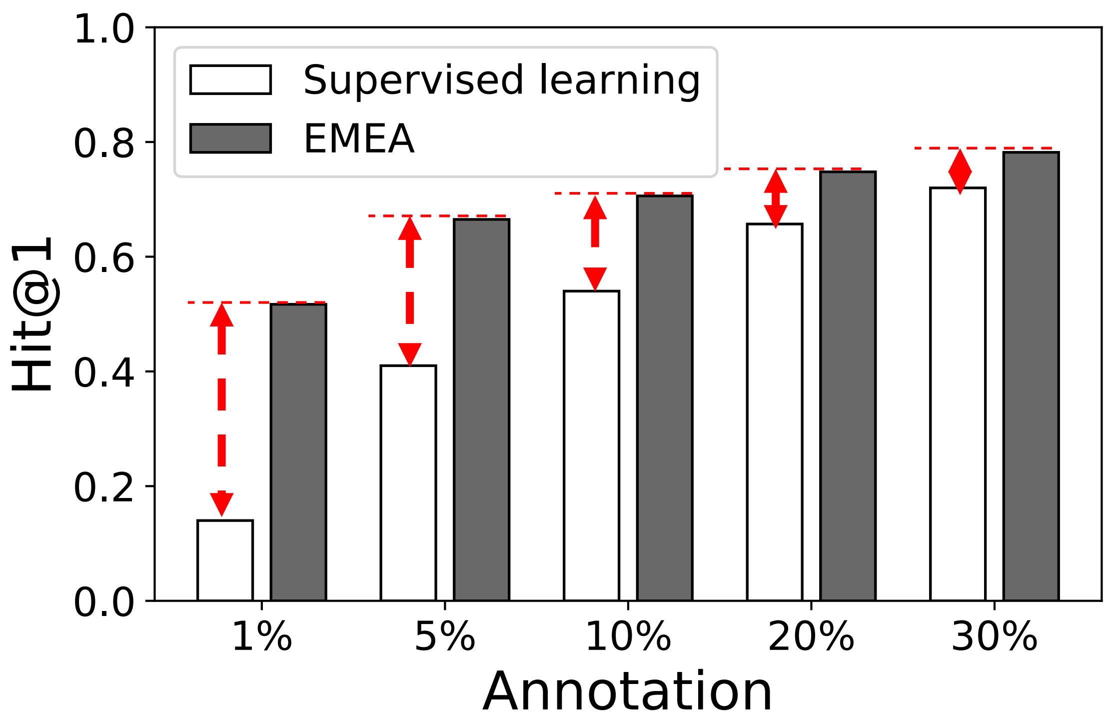


Maximization step:

- Sample $\hat{y}_U \sim q_\Theta(y_U)$
- Update Φ to maximize the $p_\Phi(y_L, y_U)$

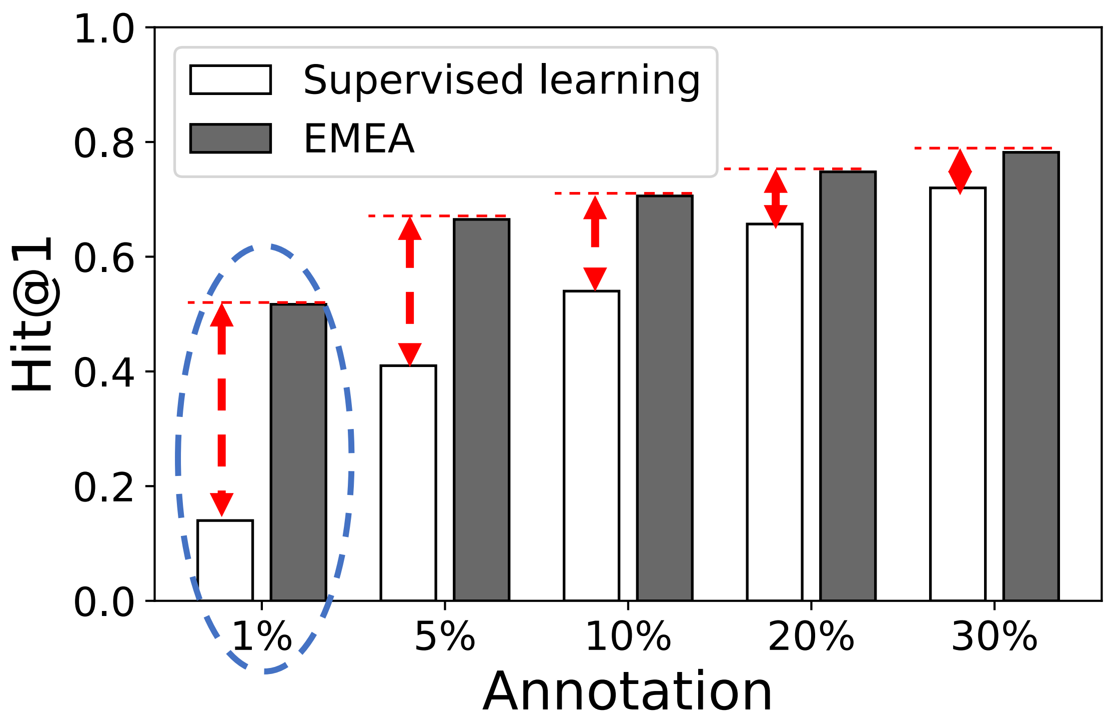
Experimental Results

Annotation Amounts



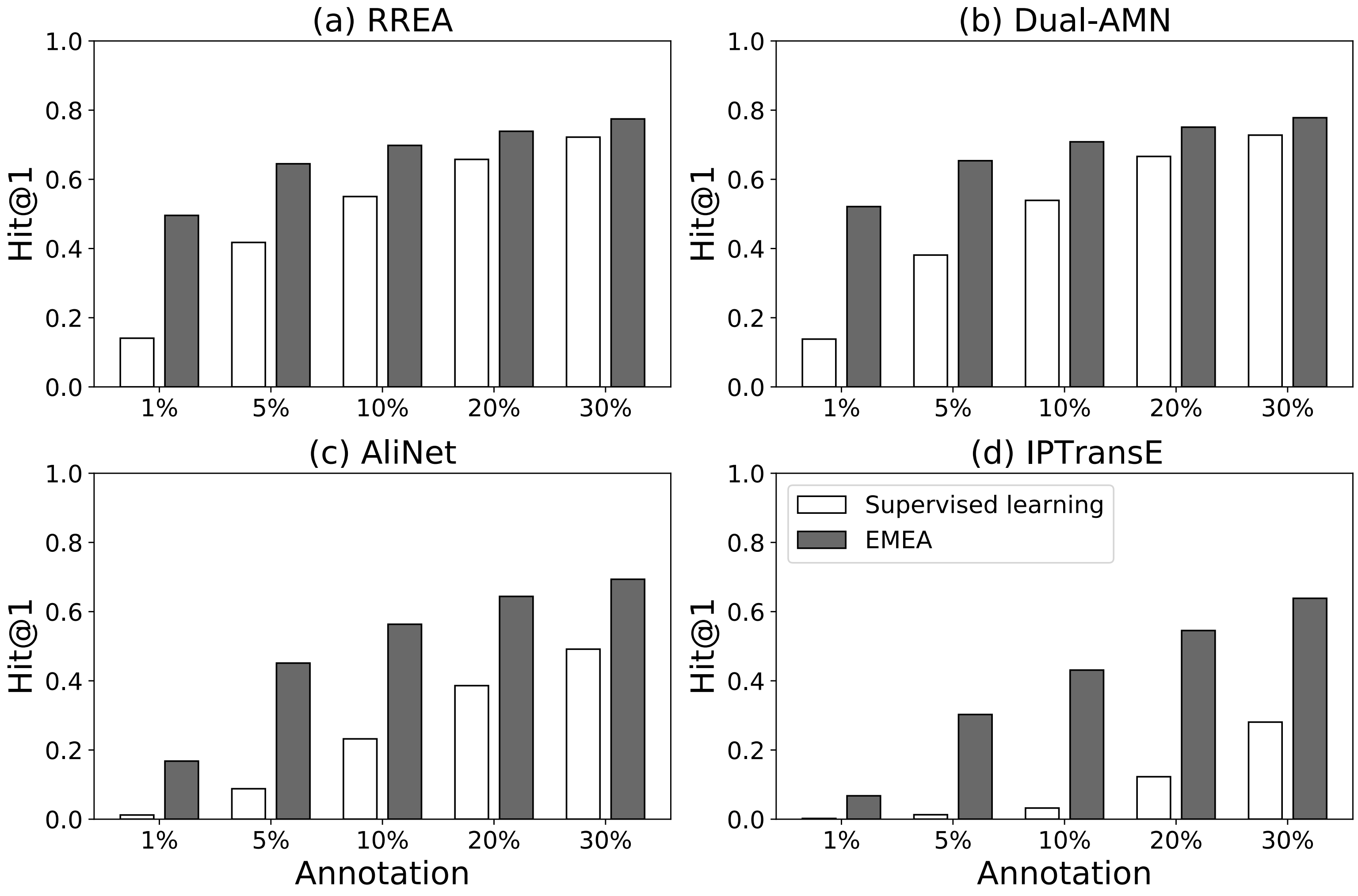
Settings: neural model: RREA, dataset: zh_en

Annotation Amounts



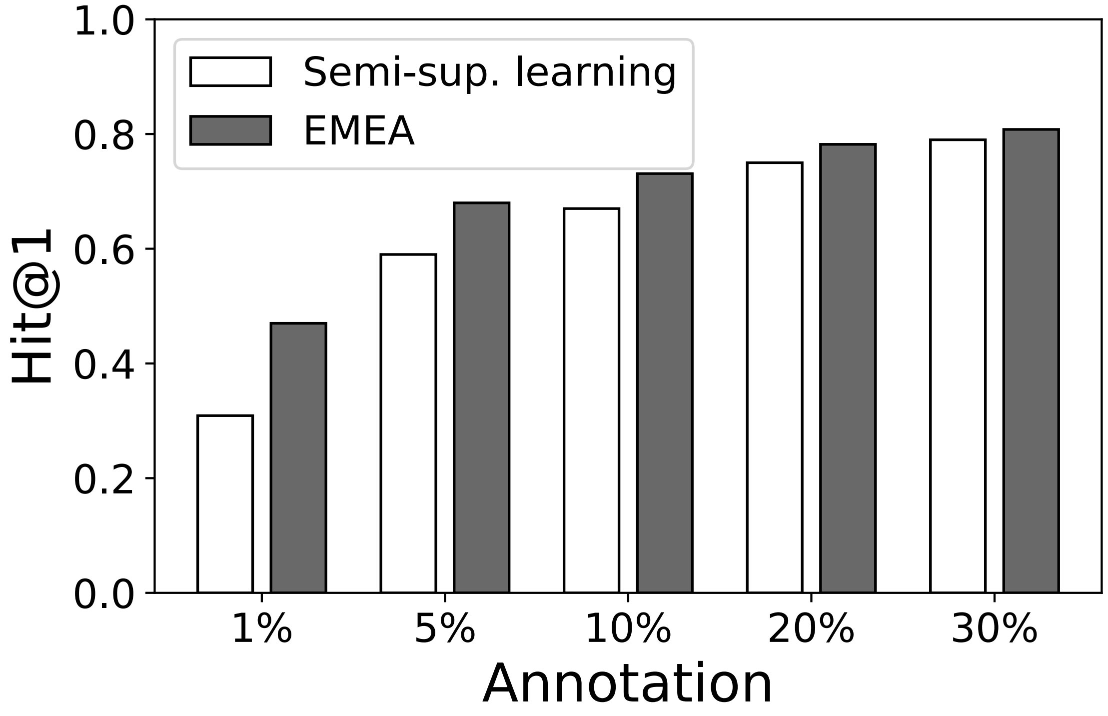
Settings: neural model: RREA, dataset: zh_en

Neural Architectures



Settings: dataset: zh_en

Initial Training Modes



Conclusion

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

EMEA: a more effective training framework of neural EA models

- Incorporate compatibility as an extra guidance of training.
- Bridge the gap between neural and reasoning-based EA methods.
- Generic across different settings.

Code & data: https://github.com/uqbingliu/EMEA

Acknowledgement

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Thank you for listening!

bing.liu@uq.edu.au

https://uqbingliu.github.io/

@BingLiu1011