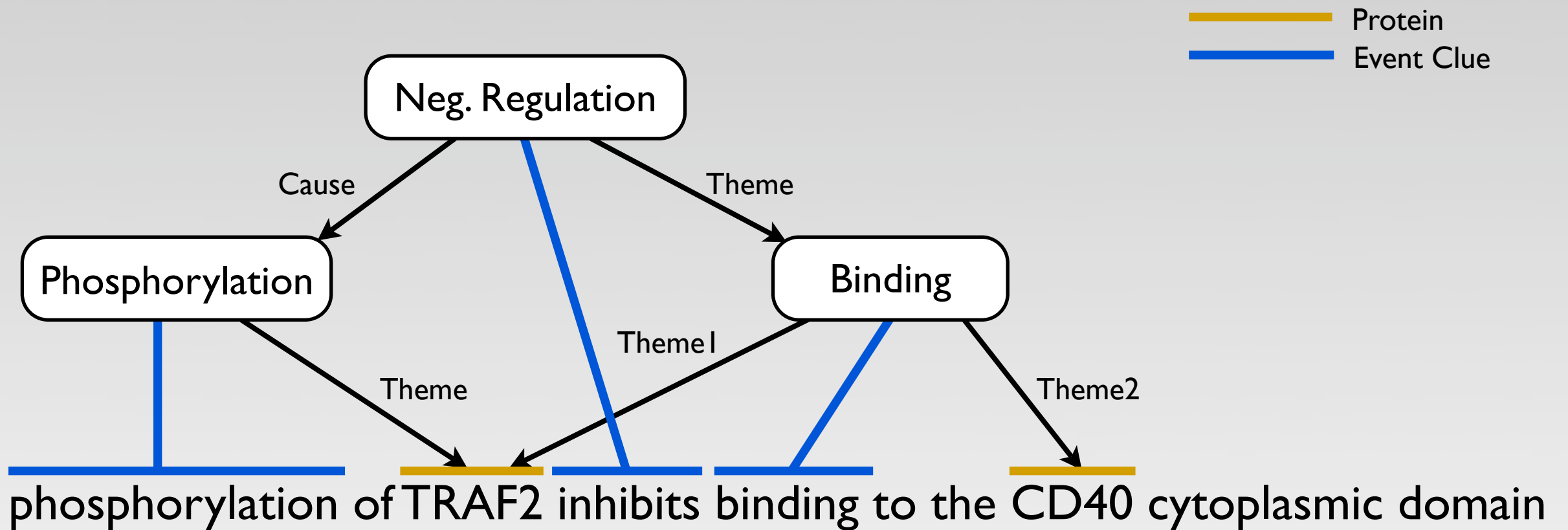# A Markov Logic Approach to Bio-Molecular Event Extraction

Sebastian Riedel, Hong-Woo Chun, Toshihisa Takagi, Jun`ichi Tsujii
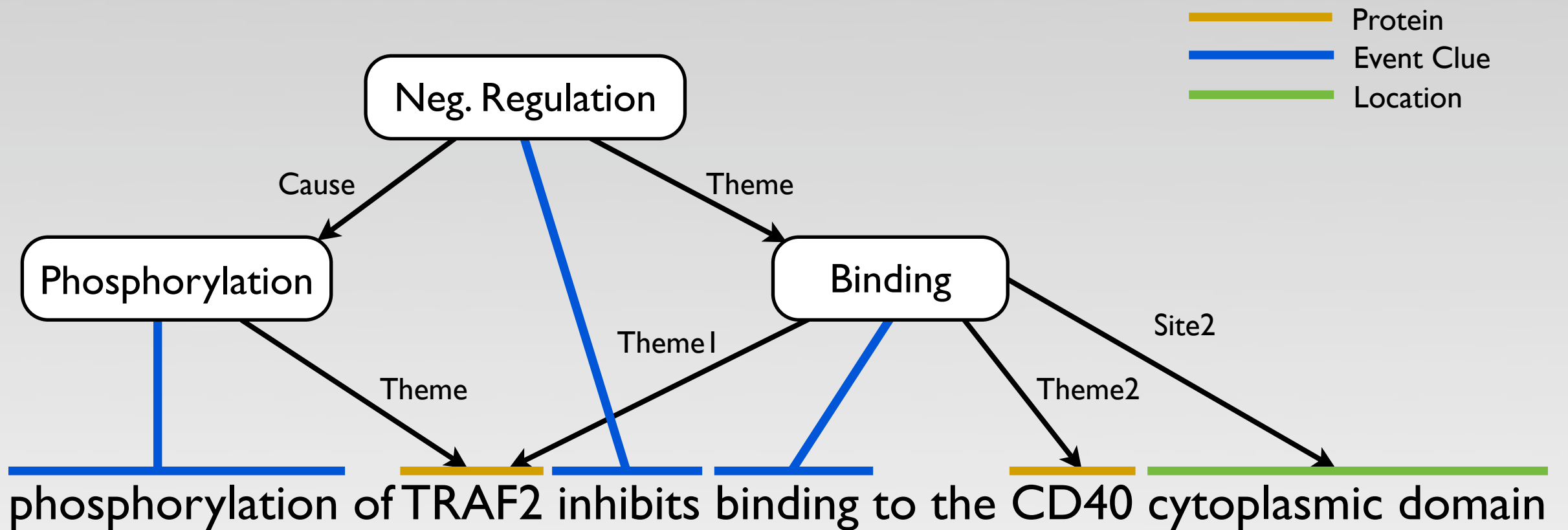
Task 1
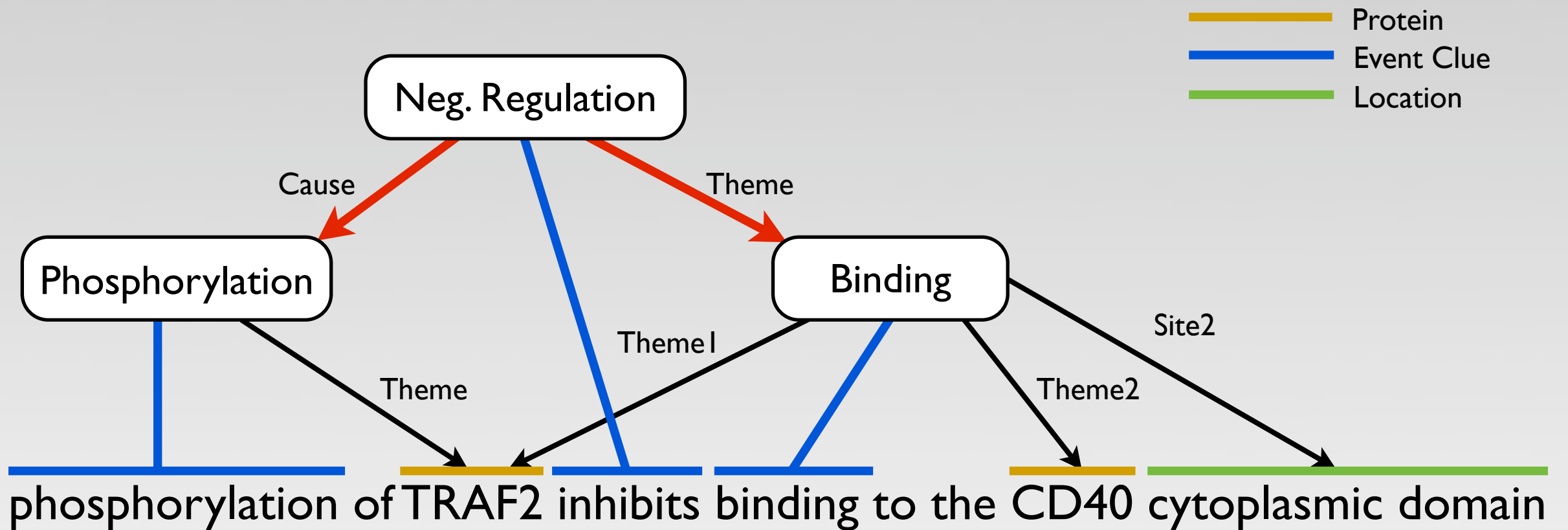
Find event **clues** and **arguments**

Task 2

- Find and attach **cellular locations**

Observations
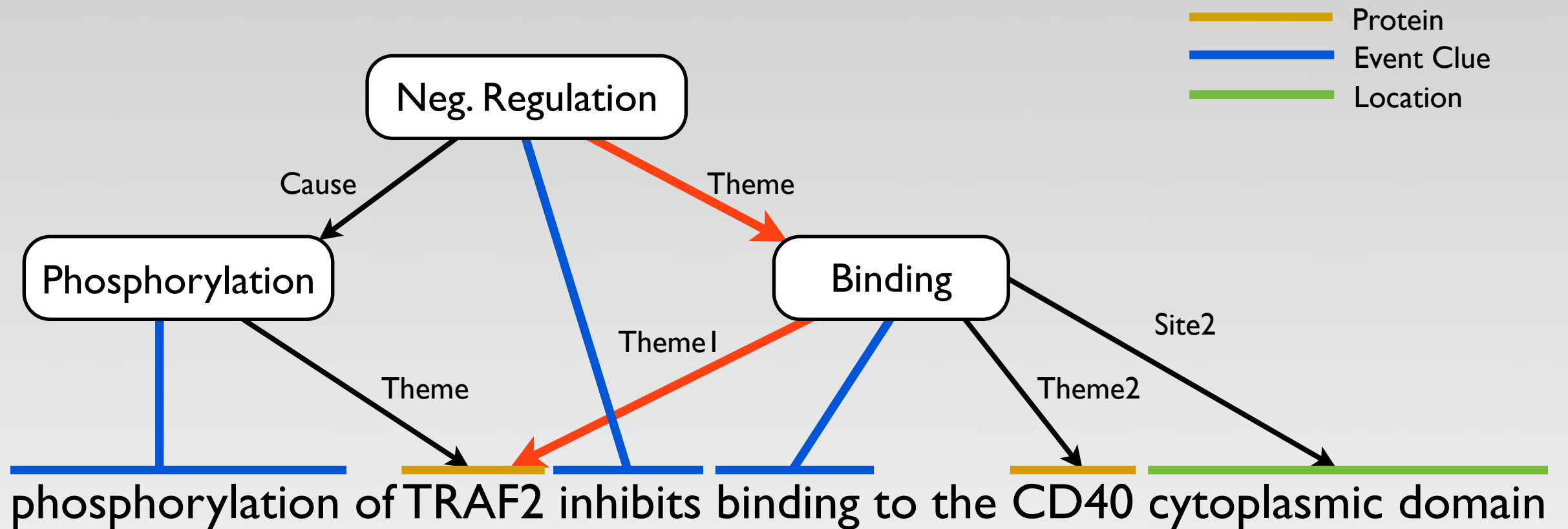
- Regulation **transitively** involves proteins

Observations
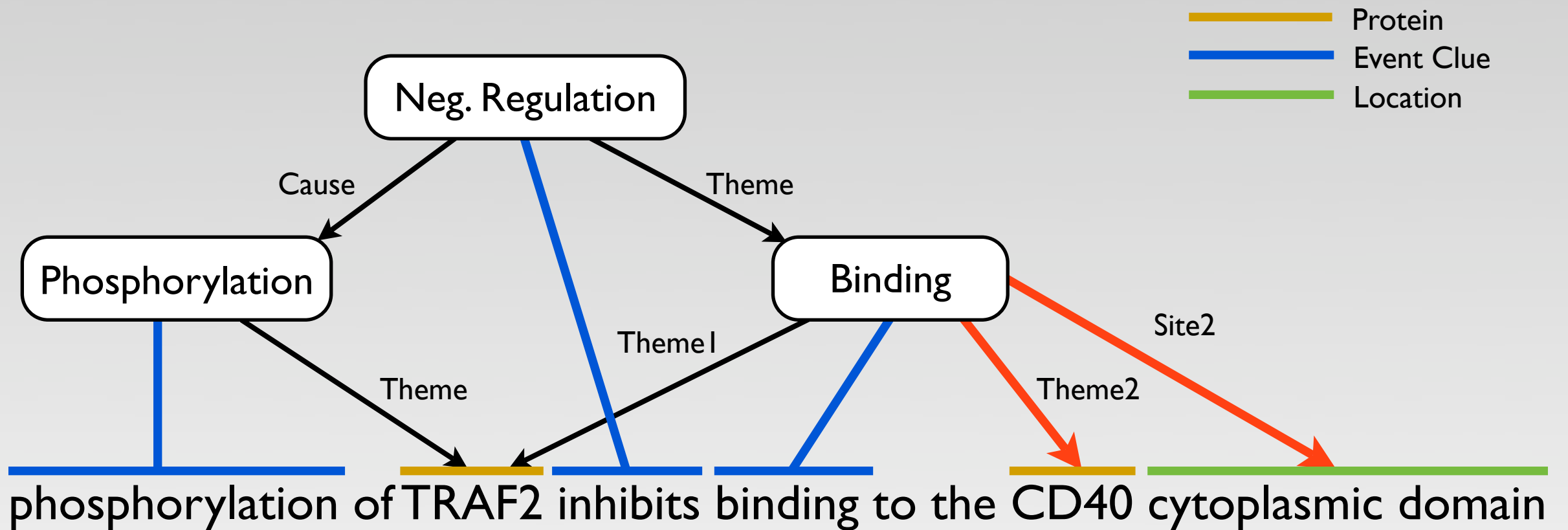
- **Site2** arguments require **Theme2**

Approach

Approach

- Due to **global interactions**:

Approach

- Due to **global interactions**:

  - Learn distribution over **full event structures**

Approach

- Due to **global interactions**:

    - Learn distribution over **full event structures**

- Joint inference is **difficult**

Approach

- Due to **global interactions**:

  - Learn distribution over **full event structures**

- Joint inference is **difficult**

  - Use **Markov Logic** & interpreter (Richardson & Domingos, 2006)

Approach

- Due to **global interactions**:

  - Learn distribution over **full event structures**

- Joint inference is **difficult**

  - Use **Markov Logic** & interpreter (Richardson & Domingos, 2006)

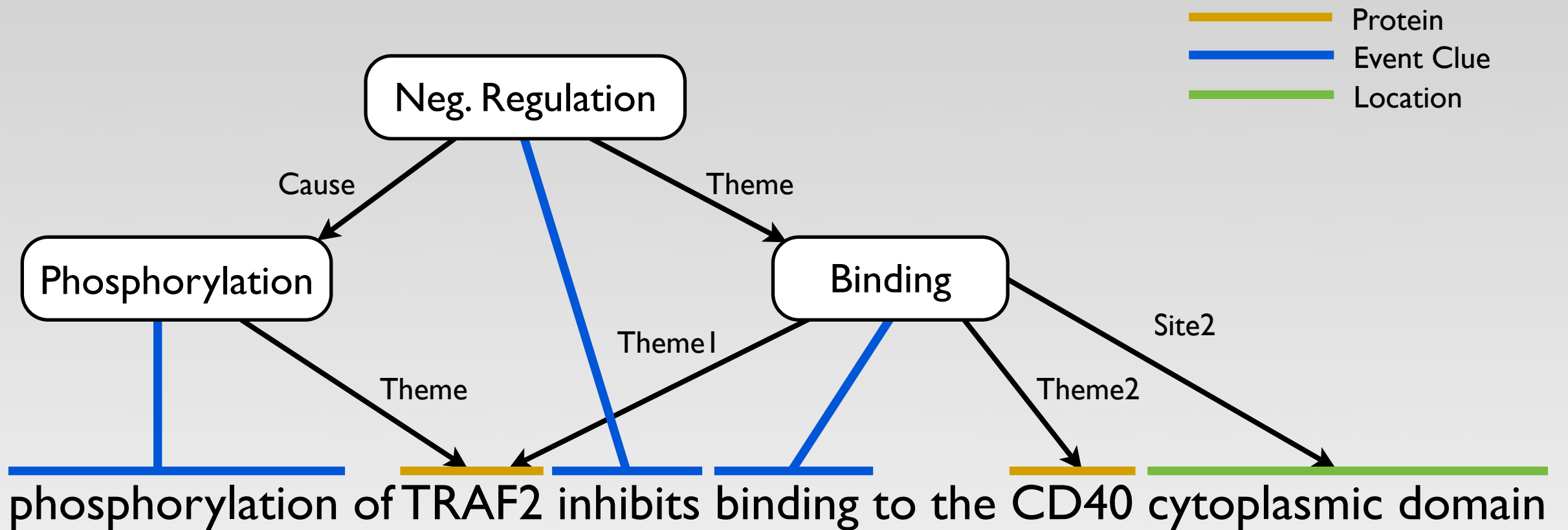- Markov Logic **likes small domains**

Approach

- Due to **global interactions**:

  - Learn distribution over **full event structures**

- Joint inference is **difficult**

  - Use **Markov Logic** & interpreter (Richardson & Domingos, 2006)

- Markov Logic **likes small domains**

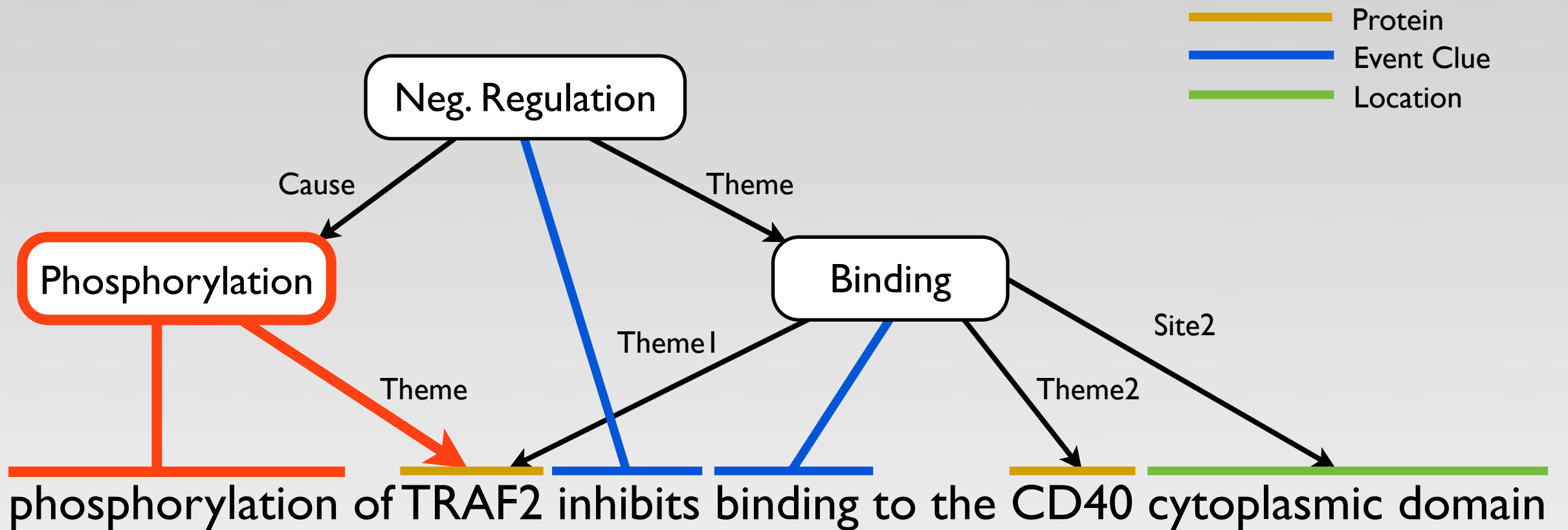 - Map to **link structure** over tokens

Event Prediction

- Event structure where events are entities

Event Prediction

- Event structure where events are entities

Link Prediction
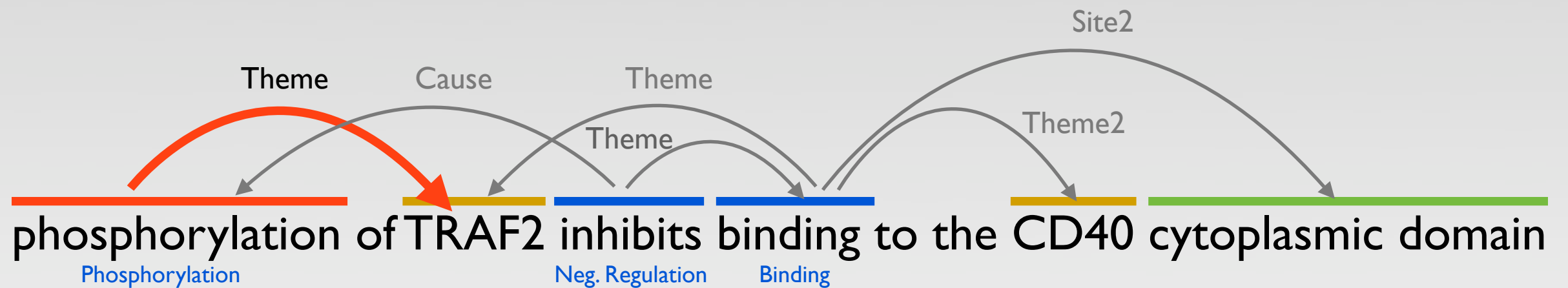
- Events are projected on their clues

Link Prediction
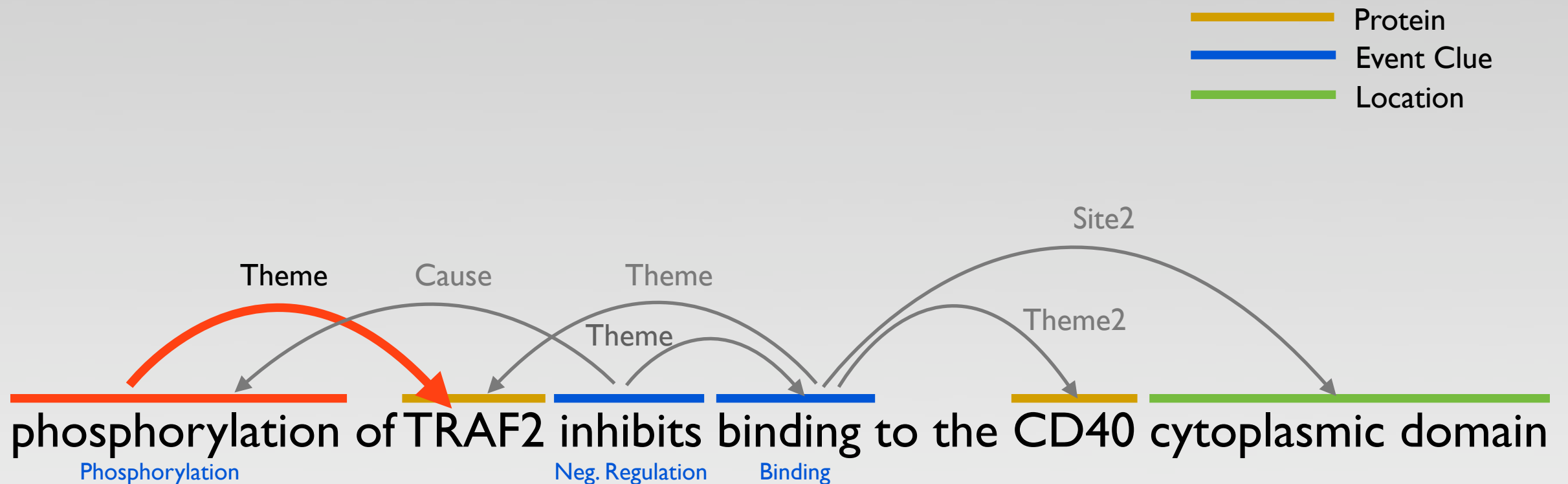
- Events are projected on their clues

- Close to **Semantic Role Labelling** : cf Meza-Ruiz & Riedel, 2009)

Markov Logic
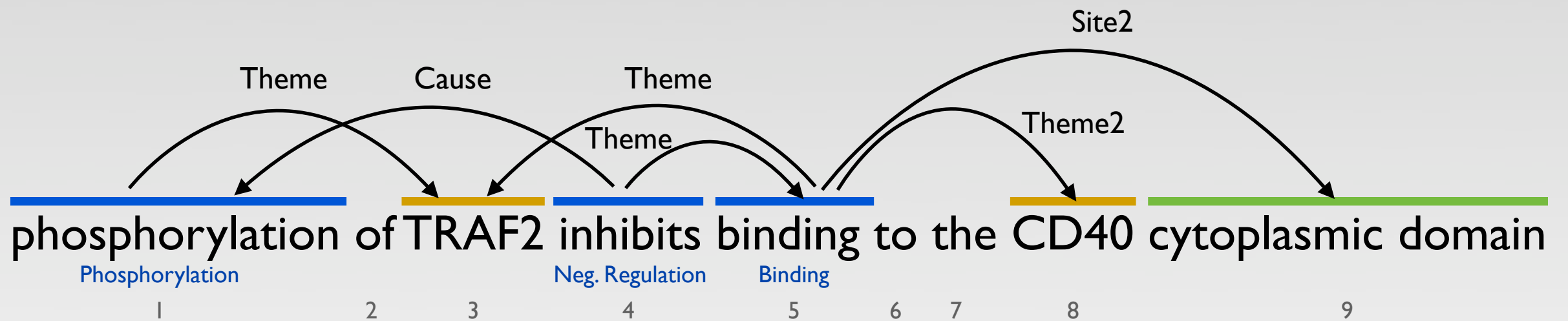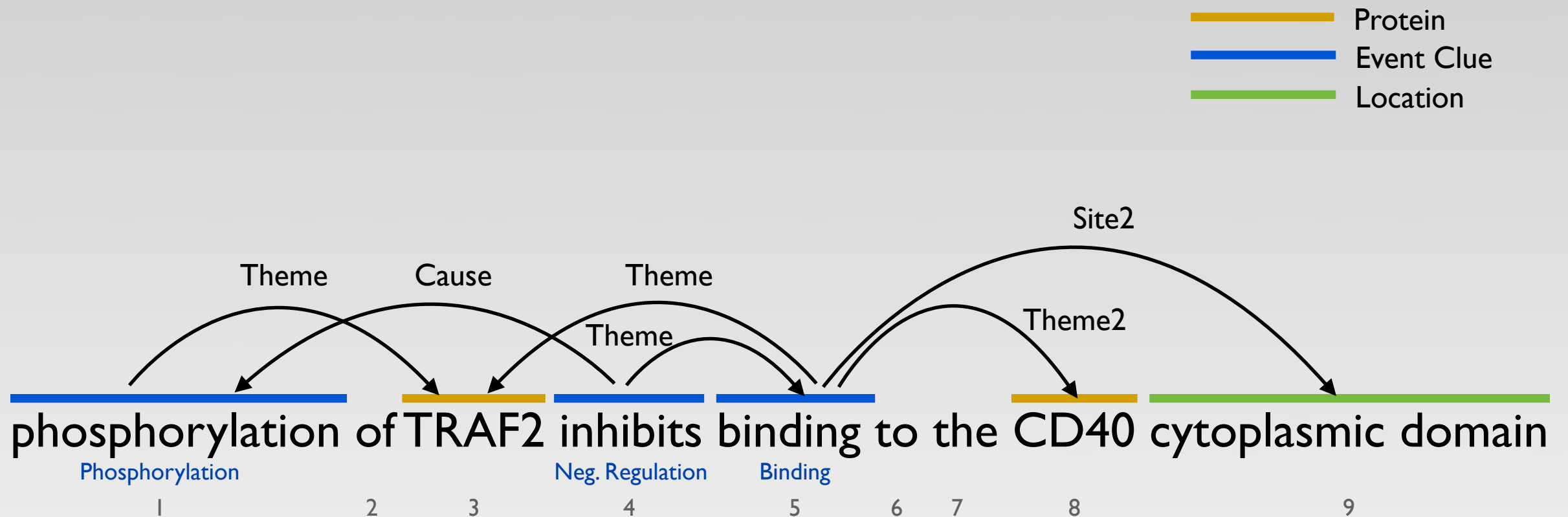
- Describe solution with **predicates** and **possible worlds**:
{*event*(1,Phosph.), *role*(1,3,Theme), *site*(9)...}
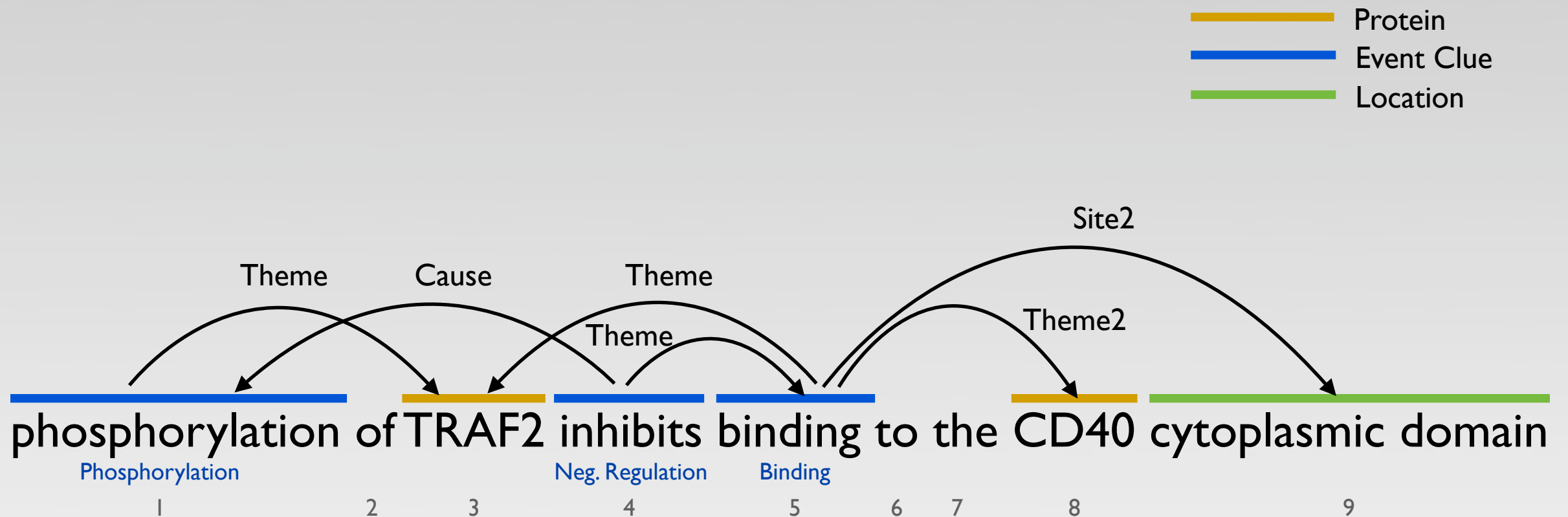
Markov Logic

- Describe solution with **predicates** and **possible worlds**:
{*event*(1,Phosph.), *role*(1,3,Theme), *site*(9)...}
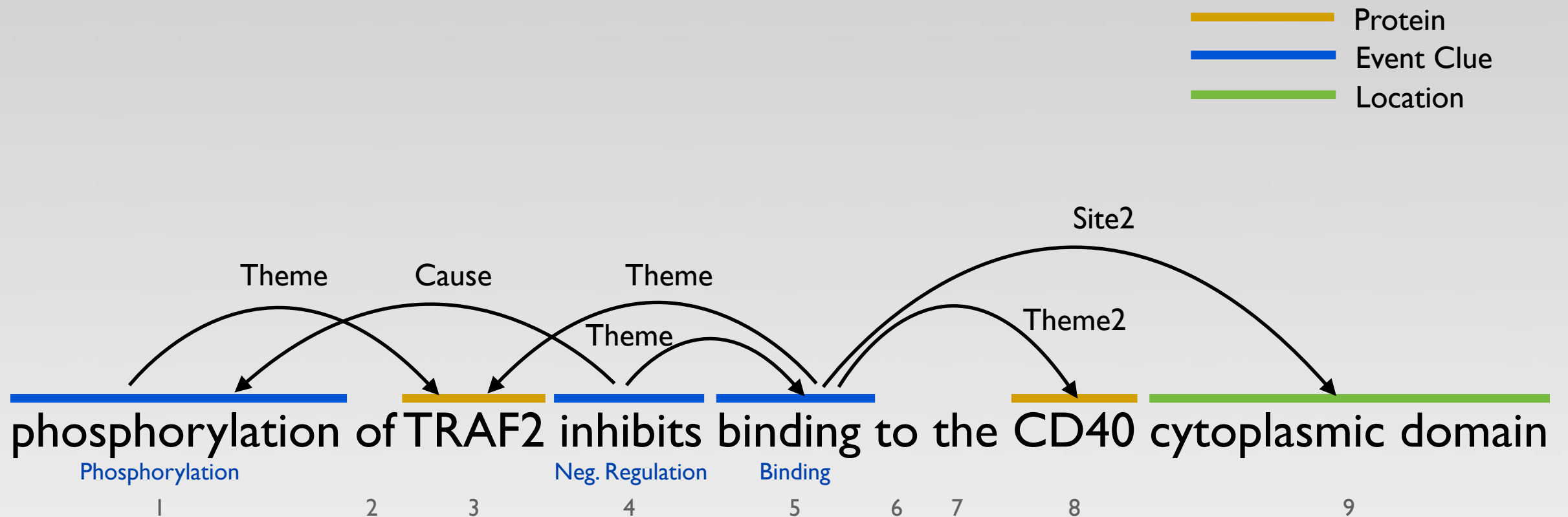
Markov Logic

- Describe solution with **predicates** and **possible worlds**:
  {*event*(1,Phosph.), *role*(1,3,Theme), *site*(9)...}

Markov Logic
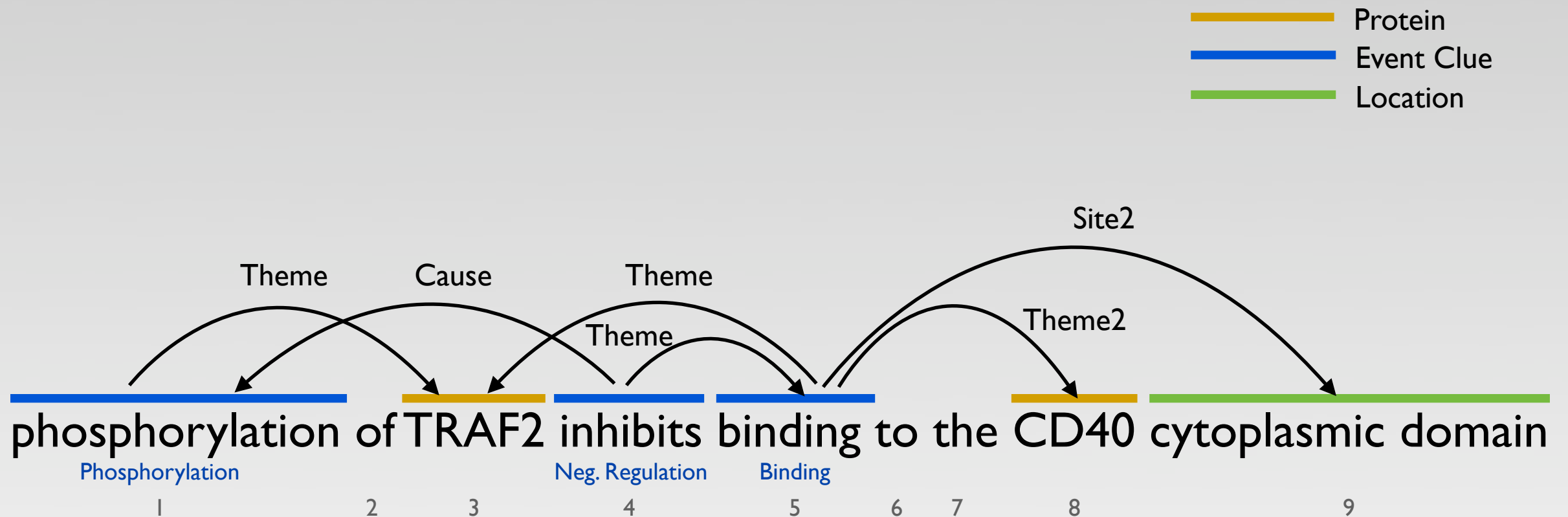
- Describe solution with **predicates** and **possible worlds**:
  {*event*(1,Phosph.), *role*(1,3,Theme), *site*(9)...}

Clue Detection

Protein
Event Clue
Location

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

Neg. Regulation

1   2   3   4   5   6 7   8   9

Clue Detection

Protein
Event Clue
Location

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

Neg. Regulation

1   2   3   4   5   6 7   8   9

- Distinguish between good and bad worlds
  word(i,inhibits) ⇒ *event(i,Neg.Regulation)* <1.2>

Site Extraction

Protein
Event Clue
Location

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

1   2   3   4   5   6 7   8   9

- **Copy and paste** formulae:
  $word(i,w) \Rightarrow location(i)$ <weight(w)>

Compact Representation

Protein
Event Clue
Location

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

Neg. Regulation

1   2   3   4   5   6 7   8   9

- Compact representation:
$word(i,w) \Rightarrow event(i,t) <weight(w,t)>$

Argument Extraction

Legend:
- Protein
- Event Clue
- Location

Theme

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

Phosphorylation — Neg. Regulation — Binding

1 2 3 4 5 6 7 8 9
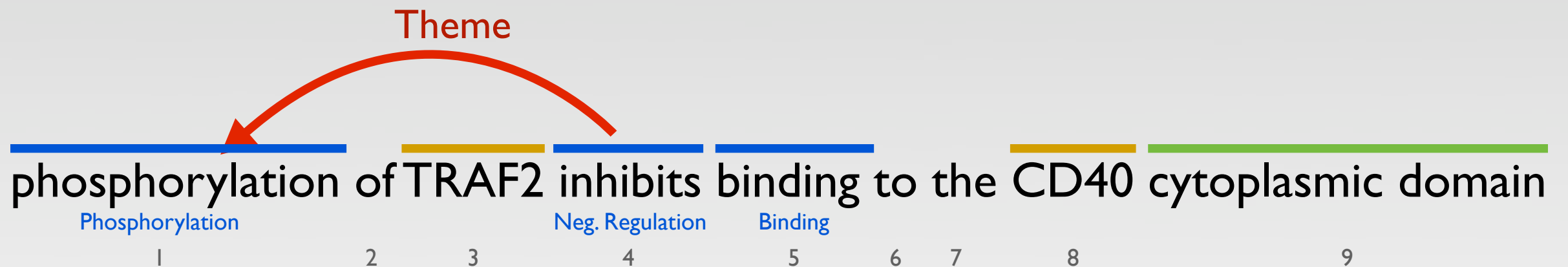
- Subjects are themes:
$dep(i,j,subj) \Rightarrow role(i,j,Theme)$ <1.23>
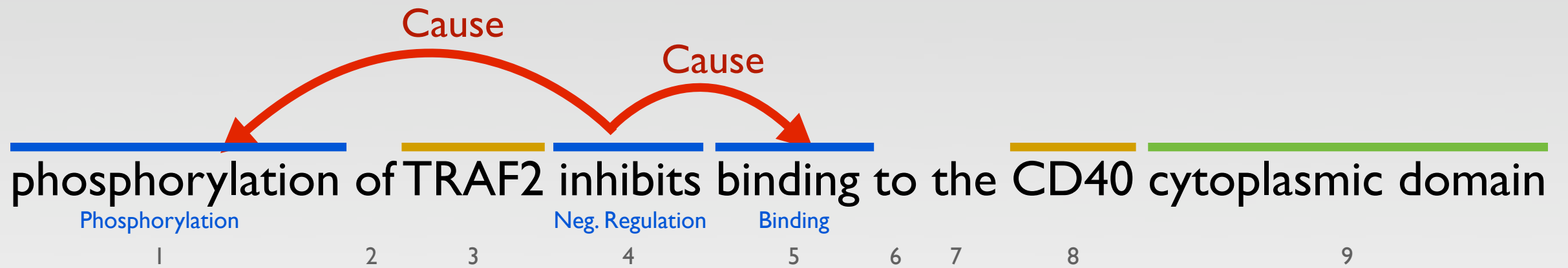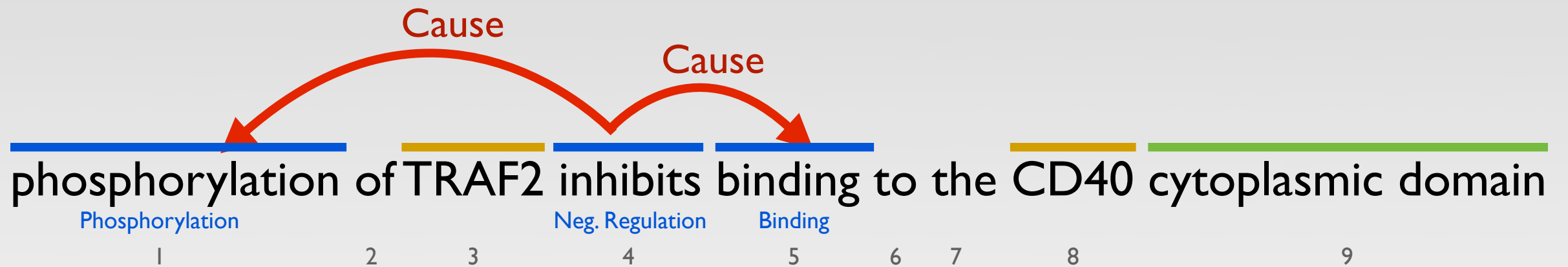
Joint Distribution

Joint Distribution

- **Ground Feature**: e.g.

$$dep(i,j,\text{subj}) \Rightarrow role(i,j,\text{Theme})$$

Joint Distribution

- **Ground Feature**: e.g.

$$dep(i,j,\text{subj}) \Rightarrow role(i,j,\text{Theme})$$

$$f^{\phi}_{j/4,i/1}(\mathbf{y}) = \begin{cases} 1 & \text{if } dep(4,1,\text{subj}) \Rightarrow role(4,1,\text{Theme}) \\ 0 & \text{otherwise} \end{cases}$$

Joint Distribution

- **Ground Feature**: e.g.

$$dep(i,j,\text{subj}) \Rightarrow role(i,j,\text{Theme})$$

$$f^{\phi}_{j/4,i/1}(\mathbf{y}) = \begin{cases} 1 & \text{if } dep(4,1,\text{subj}) \Rightarrow role(4,1,\text{Theme}) \\ 0 & \text{otherwise} \end{cases}$$

substitute free variables wrt. binding

Joint Distribution

- **Ground Feature**: e.g.

$$dep(i,j,\text{subj}) \Rightarrow role(i,j,\text{Theme})$$

$$f^{\phi}_{j/4, i/1}(\mathbf{y}) = \begin{cases} 1 & \text{if } dep\,(4, 1, \text{subj}) \Rightarrow role\,(4, 1, \text{Theme}) \\ 0 & \text{otherwise} \end{cases}$$

substitute free variables
wrt. binding

- Weighted formulae **loglinear distribution**:

Joint Distribution

- **Ground Feature**: e.g.

$$dep(i,j,\text{subj}) \Rightarrow role(i,j,\text{Theme})$$

$$f^{\phi}_{j/4,i/1}(\mathbf{y}) = \begin{cases} 1 & \text{if } dep(4,1,\text{subj}) \Rightarrow role(4,1,\text{Theme}) \\ 0 & \text{otherwise} \end{cases}$$

substitute free variables
wrt. binding

- Weighted formulae **loglinear distribution**:

$$p(\mathbf{y}) = \frac{1}{Z} \exp\left(\sum_{(\phi,w)\in M} w \sum_{\mathbf{c}\in C^{\phi}} f^{\phi}_{\mathbf{c}}(\mathbf{y}) \right)$$

Joint Distribution

- **Ground Feature**: e.g.

$dep(i,j,\text{subj}) \Rightarrow role(i,j,\text{Theme})$

$$f^{\phi}_{j/4,i/1}(\mathbf{y}) = \begin{cases} 1 & \text{if } dep\,(4,1,\text{subj}) \Rightarrow role\,(4,1,\text{Theme}) \\ 0 & \text{otherwise} \end{cases}$$

substitute free variables
wrt. binding

- Weighted formulae **loglinear distribution**:

$$p(\mathbf{y}) = \frac{1}{Z} \exp\left(\sum_{(\phi,w)\in M} w \sum_{\mathbf{c}\in C^{\phi}} f^{\phi}_{\mathbf{c}}(\mathbf{y}) \right)$$

possible world

Joint Distribution

- **Ground Feature**: e.g.

$$dep(i,j,\text{subj}) \Rightarrow role(i,j,\text{Theme})$$

$$f_{j/4,i/1}^{\phi}\left(\mathbf{y}\right) = \begin{cases} 1 & \text{if } dep\left(4,1,\text{subj}\right) \Rightarrow role\left(4,1,\text{Theme}\right) \\ 0 & \text{otherwise} \end{cases}$$

substitute free variables
wrt. binding

- Weighted formulae **loglinear distribution**:

$$p\left(\mathbf{y}\right) = \frac{1}{Z}\exp\left(\sum_{(\phi,w)\in M} w \sum_{\mathbf{c}\in C^{\phi}} f_{\mathbf{c}}^{\phi}\left(\mathbf{y}\right)\right)$$

possible world

formula

Joint Distribution

- **Ground Feature**: e.g.

$$dep(i,j,\text{subj}) \Rightarrow role(i,j,\text{Theme})$$

$$f^{\phi}_{j/4,i/1}(\mathbf{y}) = \begin{cases} 1 & \text{if } dep(4,1,\text{subj}) \Rightarrow role(4,1,\text{Theme}) \\ 0 & \text{otherwise} \end{cases}$$

substitute free variables
wrt. binding

- Weighted formulae **loglinear distribution**:

$$p(\mathbf{y}) = \frac{1}{Z} \exp\left(\sum_{(\phi,w)\in M} w \sum_{\mathbf{c}\in C^{\phi}} f^{\phi}_{\mathbf{c}}(\mathbf{y}) \right)$$

possible world

formula   weight

Joint Distribution

- **Ground Feature**: e.g.

$$dep(i,j,\text{subj}) \Rightarrow role(i,j,\text{Theme})$$

$$f^{\phi}_{j/4,i/1}(\mathbf{y}) = \begin{cases} 1 & \text{if } dep\,(4,1,\text{subj}) \Rightarrow role\,(4,1,\text{Theme}) \\ 0 & \text{otherwise} \end{cases}$$

substitute free variables
wrt. binding

- Weighted formulae **loglinear distribution**:

ground feature

$$p(\mathbf{y}) = \frac{1}{Z} \exp\left(\sum_{(\phi,w) \in M} w \sum_{\mathbf{c} \in C^{\phi}} f^{\phi}_{\mathbf{c}}(\mathbf{y}) \right)$$

possible world

formula   weight

Joint Distribution

- **Ground Feature**: e.g.

$$dep(i,j,\text{subj}) \Rightarrow role(i,j,\text{Theme})$$

$$f_{j/4,i/1}^{\phi}(\mathbf{y}) = \begin{cases} 1 & \text{if } dep\,(4,1,\text{subj}) \Rightarrow role\,(4,1,\text{Theme}) \\ 0 & \text{otherwise} \end{cases}$$

substitute free variables
wrt. binding

- Weighted formulae **loglinear distribution**:

ground feature

$$p(\mathbf{y}) = \frac{1}{Z} \exp \left(\sum_{(\phi,w)\in M} w \sum_{\mathbf{c}\in C^{\phi}} f_{\mathbf{c}}^{\phi}(\mathbf{y}) \right)$$

possible world

formula   weight   binding of free variables

Learning and Inference

Learning and Inference

- How do we find most likely structure?

Learning and Inference

- How do we find most likely structure?

  - Cutting Plane Inference+ILP (Riedel, 2008)

Learning and Inference

- How do we find most likely structure?

  - Cutting Plane Inference+ILP (Riedel, 2008)

- How do we learn weights?

Learning and Inference

- How do we find most likely structure?

  - Cutting Plane Inference+ILP (Riedel, 2008)

- How do we learn weights?

  - Single best MIRA (Krammer, 2006)

System

System

- Local Formulae

System

- Local Formulae

  - Dependency paths (labelled, unlabelled)

System

- Local Formulae
  - Dependency paths (labelled, unlabelled)
  - Words, POS tags, Stems, 2 Dictionaries

System

- Local Formulae

  - Dependency paths (labelled, unlabelled)

  - Words, POS tags, Stems, 2 Dictionaries

- Global Formulae

System

- Local Formulae

  - Dependency paths (labelled, unlabelled)

  - Words, POS tags, Stems, 2 Dictionaries

- Global Formulae

  - (Inspired by the shared task validator)

Results

Results

- **44.4** F-score for Task 1 (4th rank)

Results

- **44.4** F-score for Task 1 (4th rank)

- Using dev set for training: 45.1 (3th rank)

Results

- **44.4** F-score for Task 1 (4th rank)

- Using dev set for training: 45.1 (3th rank)

- **43.1** F-score for Task 2 (1st rank)

Results

- **44.4** F-score for Task 1 (4th rank)

- Using dev set for training: 45.1 (3th rank)

- **43.1** F-score for Task 2 (1st rank)

- Poor results for **Bindings**

Results

- **44.4** F-score for Task 1 (4th rank)

- Using dev set for training: 45.1 (3th rank)

- **43.1** F-score for Task 2 (1st rank)

- Poor results for **Bindings**

- Fast to train: **3 hours** on MacBook Pro

Impact of Joint Inference?

Impact of Joint Inference?

- Formula sets:

Impact of Joint Inference?

- Formula sets:

  - **CORE**: not more than one type, role

Impact of Joint Inference?

- Formula sets:

  - **CORE**: not more than one type, role

  - **VALID**: consistent arguments and events

Impact of Joint Inference?

- Formula sets:

    - **CORE**: not more than one type, role

    - **VALID**: consistent arguments and events

    - **FULL**: "additional" formulae

Impact of Joint Inference (on Atom F1)

Predicate	CORE	VALID	FULL
total	50.7	60.1	61.9
event	52.8	63.2	64.3
role	44.0	53.5	55.7
site	42.0	46.0	51.5

Impact of Joint Inference (on Atom F1)

Predicate	CORE	VALID	FULL
total	50.7	60.1	61.9
event	52.8	63.2	64.3
role	44.0	53.5	55.7
site	42.0	46.0	51.5

- FULL does not explicitly consider *site*

- But Turku still did better w/ a local model

Conclusion

Conclusion

- **Joint** and **Declarative** Approach

Conclusion

- **Joint** and **Declarative** Approach

- Global Formulae help **across the board**

Conclusion

- **Joint** and **Declarative** Approach

- Global Formulae help **across the board**

- Location extraction **easy to implement**

Conclusion

- **Joint** and **Declarative** Approach

- Global Formulae help **across the board**

- Location extraction **easy to implement**

- **Compact** representation of model