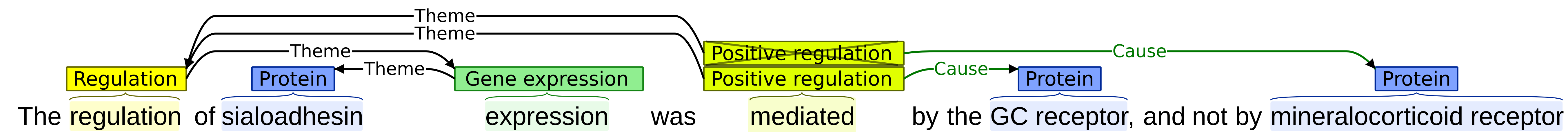


# New Resources and Perspectives for Biomedical Event Extraction

## Event extraction (EE)

- Information extraction using expressive structured representations
- Major focus of recent work in BioNLP: BioNLP Shared Tasks (ST) in 2009 and 2011
- Many systems and resources introduced
- Despite progress, many challenges remain



**This work**

- Compilation of analyses from broad set of EE systems
- Analysis of events that cannot be recovered by any system
- Document-level representation and evaluation for EE
- Document-level evaluation of BioNLP ST'11 systems
- All newly introduced resources and tools available

## New Resources

### Compilation of automatically created analyses from BioNLP Shared Task 2011 event extraction systems

Team	BioNLP ST'11 tasks	Contributors / system description authors (proc. BioNLP ST'11)
UTurku	(all eight tasks)	Björne and Salakoski (2011)
ConcordU	GE, EPI, ID, CO, REL, REN	Kilicoglu and Bergler (2011)
UMass	GE, EPI, ID	Riedel and McCallum (2011)
Stanford	GE, EPI, ID	McClosky, Surdeanu and Manning (2011)
FAUST	GE, EPI, ID	Riedel, McClosky, Surdeanu, McCallum and Manning (2011)
MSR-NLP	GE, EPI	Quirk, Choudhury, Gamon and Vanderwende (2011)
CCP-BTMG	GE, EPI	Liu, Komandur, and Verspoor (2011)
BMI@ASU	GE	Emadzadeh, Nikfarjam and Gonzalez (2011)
TM-SCS	GE	Bui and Sloot (2011)
UWMadison	GE	Vlachos and Craven (2011)
HCMUS	GE, REL	Le Minh, Nguyen Truong, and Ho Bao (2011)
PredX	ID	Gobert and Royauté
VIBGhent	REL	Van Landeghem, Abeel, De Baets and Van de Peer (2011)

Both final test set submissions and representative development set submissions for all participants included

### Maximum recall analysis

Task	Recall
GE	69%
EPI	79%
IE	74%

Results for union of systems on development sets of three ST'11 main tasks (task 1/core)

All resources available:  
<http://2011.bionlp-st.org>

### Common properties of events not extracted by any system

Event property	Count
Implicit argument	66
Cross-sentence	58
Weak trigger	53
Coreference	50
Static Relation	40
Error in gold	30
Ambiguous type	22
Shared trigger	15

Results for manual analysis of properties of events not extracted by any system.

## New Perspectives

### BioNLP Shared Task

#### Text-bound annotation

- each associated with a specific span of text

#### Instance-based evaluation

- all annotations considered independently

Detailed evaluation and system output, but may not reflect real-world extraction needs, and excludes some extraction approaches

### Proposed approach

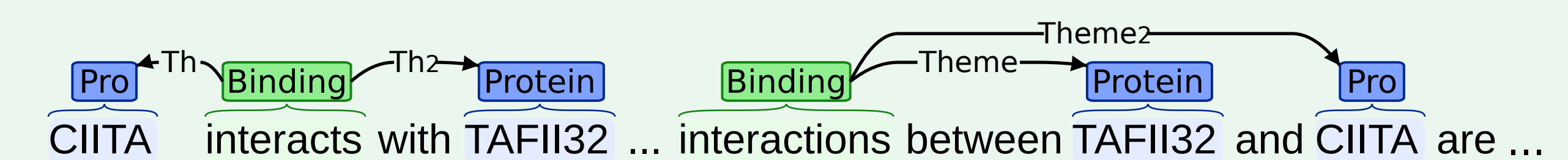
#### Document-level annotation

- no requirement to identify "trigger" texts

#### "Off the page" unique event-based evaluation

- identical events need only be extracted once

Less detailed system output, but fewer constraints on extraction approaches, evaluation more accurately reflects needs of applications such as semantic search



### BioNLP Shared Task

T1	Protein	0	5	CIITA
T2	Protein	21	28	TAFII32
T3	Binding	6	15	interacts
E1	Binding:T3	Theme:T1	Theme2:T2	
T4	Protein	54	61	TAFII32
T5	Protein	66	71	CIITA
T6	Binding	33	45	interactions
E2	Binding:T6	Theme:T4	Theme2:T5	

### Document level

T1	Protein	CIITA
T2	Protein	TAFII32
E1	Binding	Theme:T1 Theme2:T2

### Comparison of results for instance-level and document-level evaluation

