Explainable Temporal Fact Validation Through Temporal Constraints Discovery In Knowledge Graphs

FATIHA SAÏS

Joint Work With:

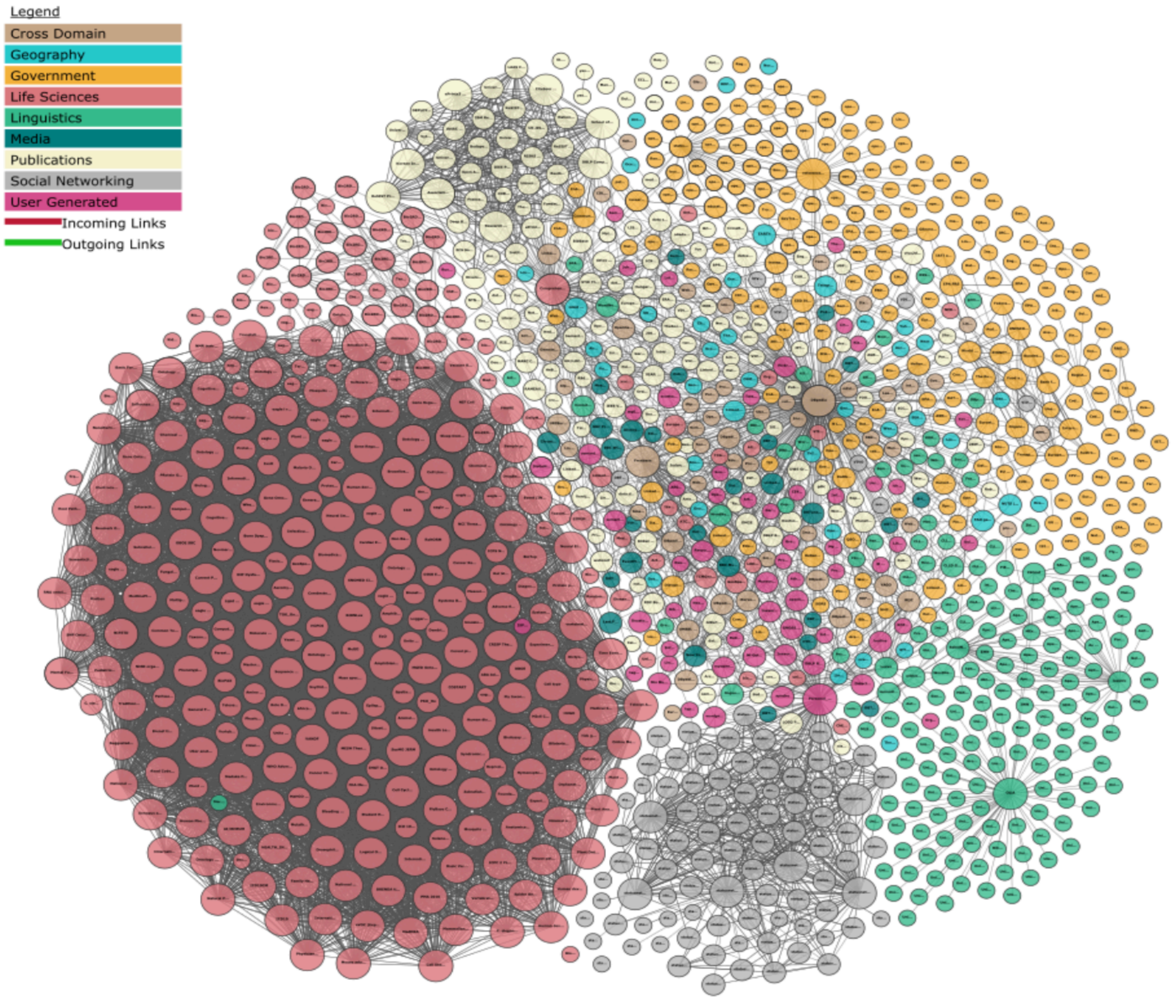
T. Soulard, J. Raad,

J. E. MALAVERRI, G. QUERCINI

LISN, CNRS & Université Paris Saclay

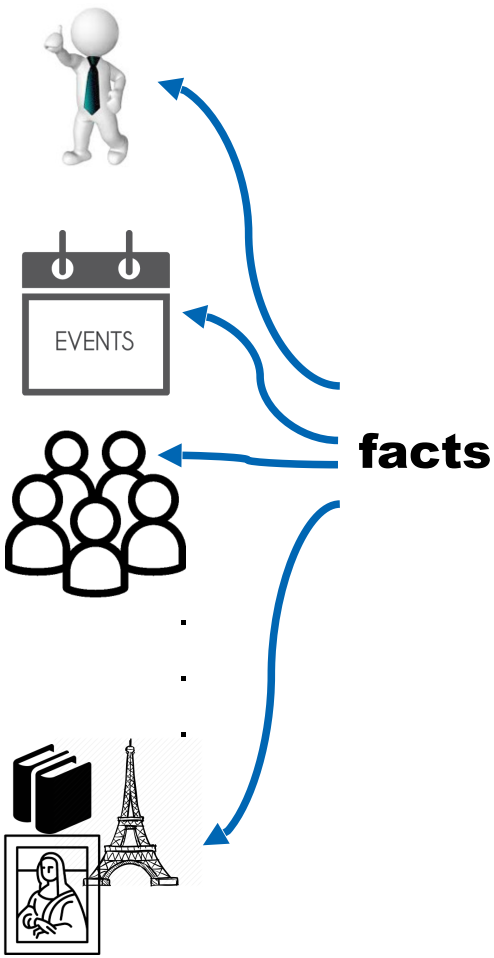
Séminaire SESAME - INRAE - 10/06/2024

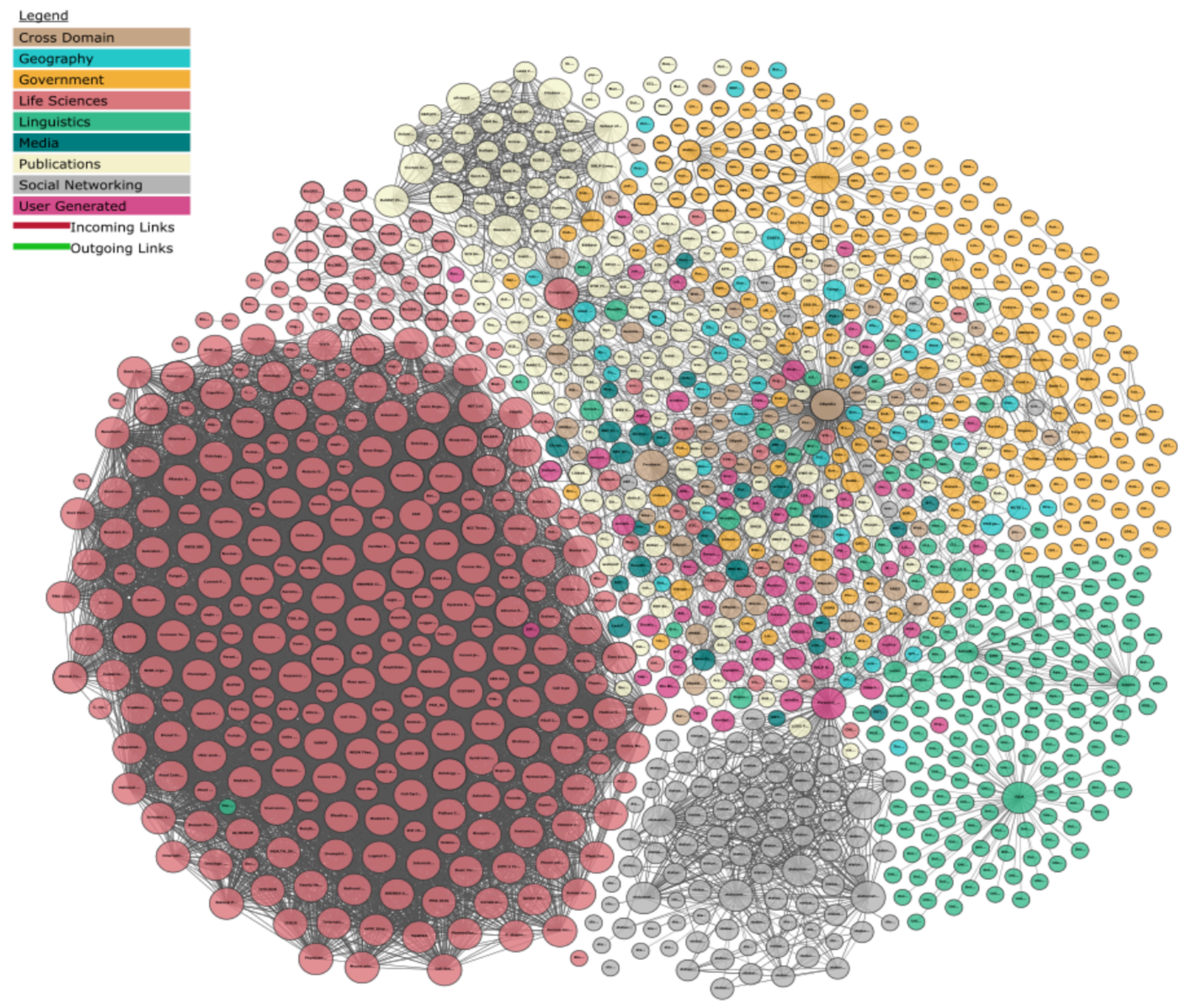
KNOWLEDGE GRAPHS (KGS)



"Linking Open Data cloud diagram 2017, by Andrejs Abele, John P. McCrae, Paul Buitelaar, Anja Jentzsch and Richard Cyganiak. http://lod-cloud.net/"

KNOWLEDGE GRAPHS (KGS)



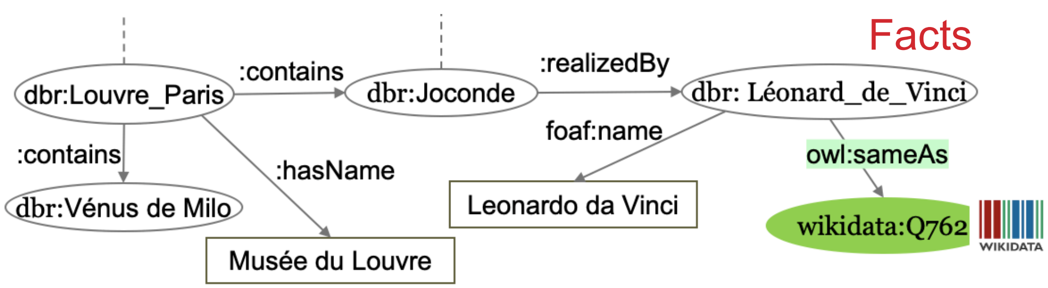


"Linking Open Data cloud diagram 2017, by Andrejs Abele, John P. McCrae, Paul Buitelaar, Anja Jentzsch and Richard Cyganiak. http://lod-cloud.net/"

KNOWLEDGE GRAPHS (KGS)

Knowledge representation model

RDF graphs: facts (data) describing entities

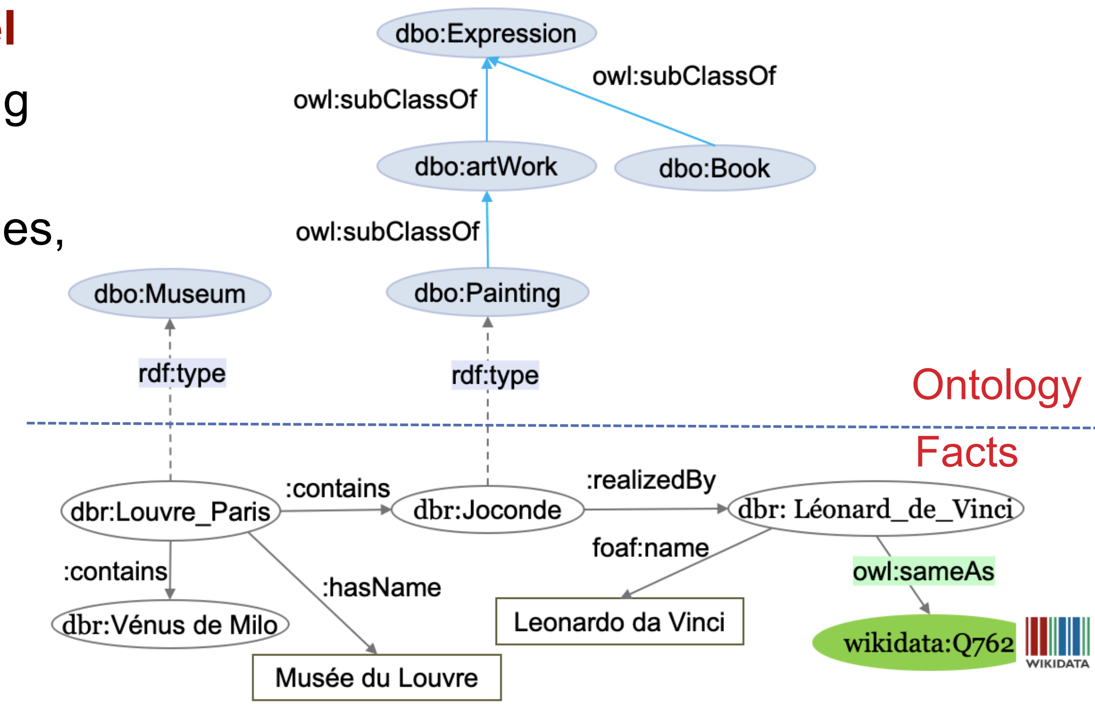


KNOWLEDGE GRAPHS (KGS)

Knowledge representation model

RDF graphs: facts (data) describing entities

OWL Ontology: concepts, properties, axioms/constraints of the domain

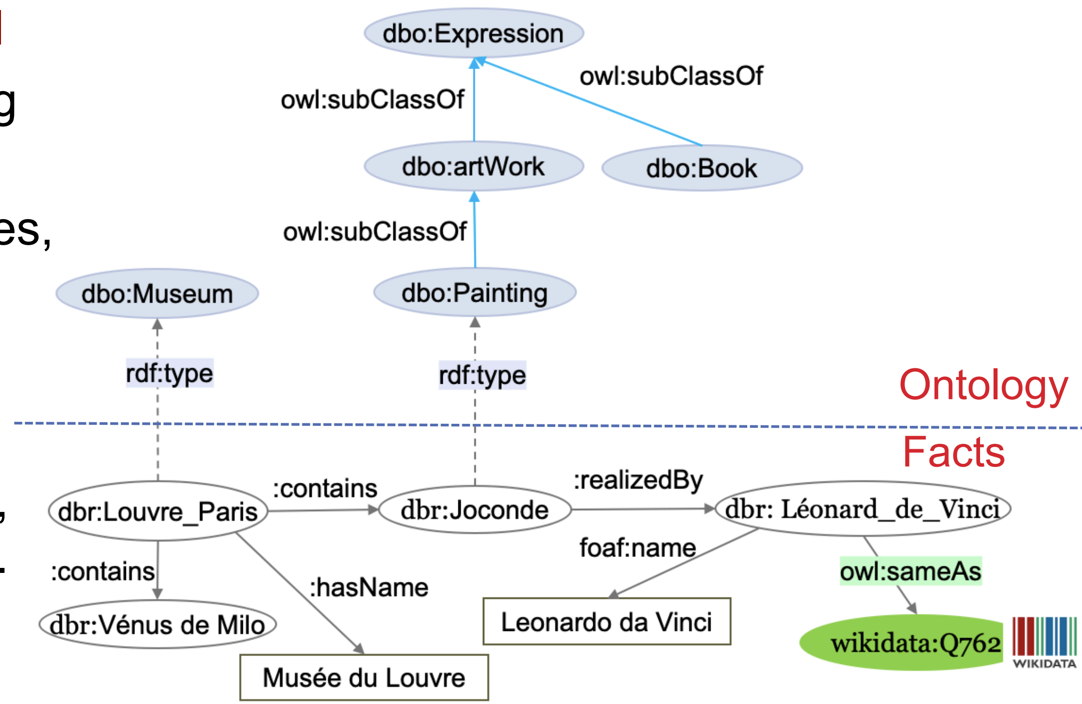


KNOWLEDGE GRAPHS (KGS)

Knowledge representation model

RDF graphs: facts (data) describing entities

OWL Ontology: concepts, properties, axioms/constraints of the domain



Applications

Web search, conversational agents, recommendation, transparency, etc.

Challenges

Enriching knowledge graphs

Guaranteeing the validity of data and knowledge

2007. 2012 …

KNOWLEDGE GRAPHS: SOME KEY PROBLEMS

- **KG Creation**
  - Information extraction from web pages (DBpedia, Yago), collaborative (Wikidata), ...
- **KG Enrichment and Expansion**
  - Link prediction, data fusion, data/knowledge linking, ...

.

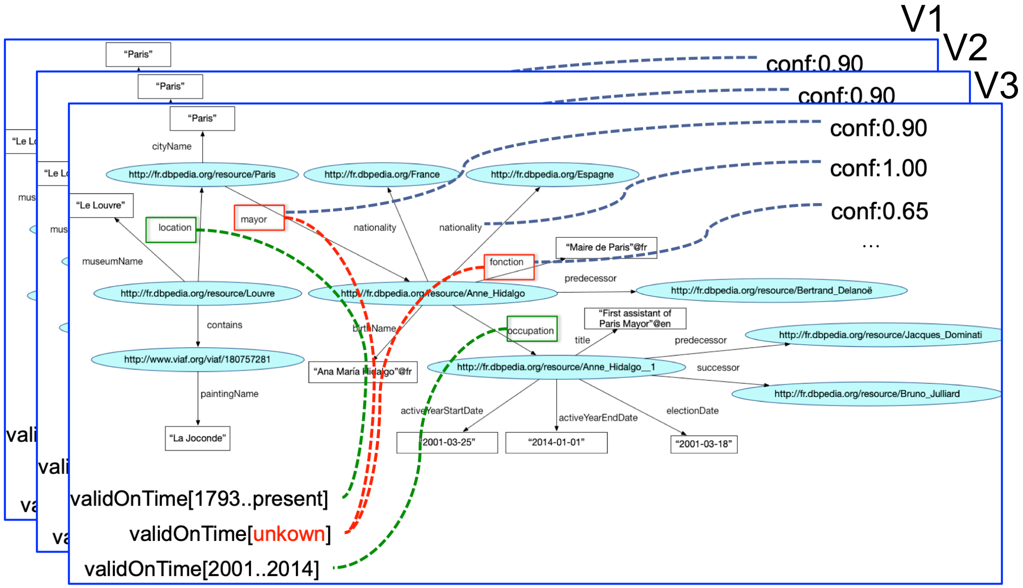
KNOWLEDGE GRAPHS: SOME KEY PROBLEMS

- **KG Creation**
  - Information extraction from web pages (DBpedia, Yago), collaborative (Wikidata), ...
- **KG Enrichment and Expansion**
  - Link prediction, data fusion, data/knowledge linking, ...
- **KG Validity**
  - Dealing with errors and ambiguities
  - Fact validity
  - Timeliness: truth of statements often changes with time:
    - E.g., Currently, who is the US president?

KGS – WHY TIME MATTERS?

Need to capture and reason on temporal information

- KG content cannot be assumed to be static, because many facts change over time
- Ignoring such temporal information may lead to ambiguity and misunderstanding



KGS – WHY TIME MATTERS?

Need to capture and reason on temporal information

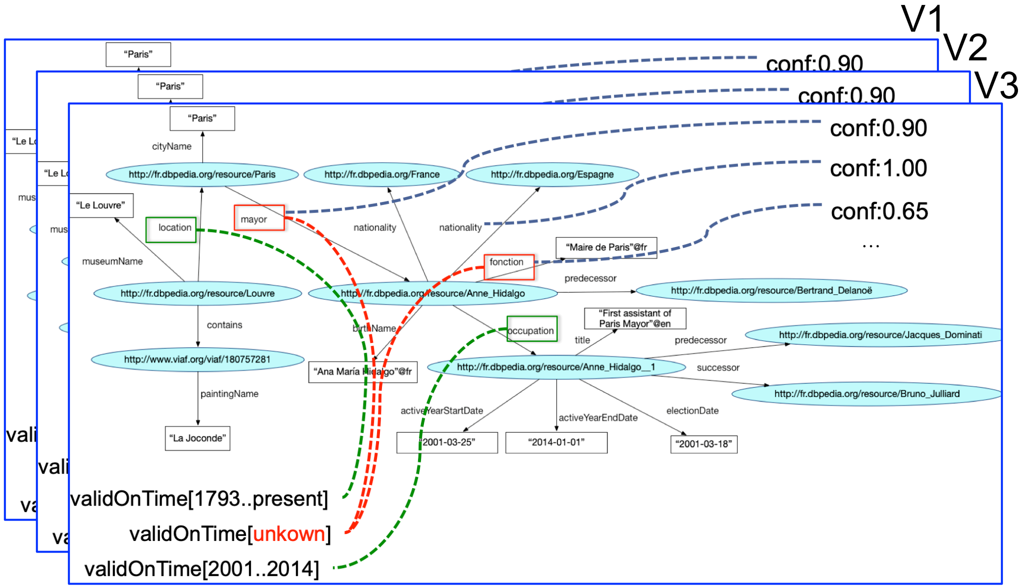
- KG content cannot be assumed to be static, because many facts change over time
- Ignoring such temporal information may lead to ambiguity and misunderstanding

Time period during which a fact is valid: validity time.

- f1: <#Obama presidentOf US> is true in the temporal context [2008-2016]
- f2: <#Trump presidentOf US> is true in the temporal context [2017-2021]

Important for :

- Query answering,
- Consistency checking,
- Knowledge discovery, …



TWO MAIN PROBLEMS

- (A) How can we generate temporal meta-facts using only the facts and the structure of the KGs?

[Malaverri et al. 2020]

- (B) How can we assess the temporal validity of facts ?

[Soulard et al. 2024]

(A) TEMPORAL META-FACT EXTRACTION - PROPOSED APPROACH

[Malaverri et al. 2020]

- How can we generate temporal meta-facts using only the facts and the structure of the KGs?

Our approach

1. Generate seed meta-facts
2. Use the KG structure to propagate temporal meta-facts
3. Exploit these temporal meta-facts to asses facts veracity.

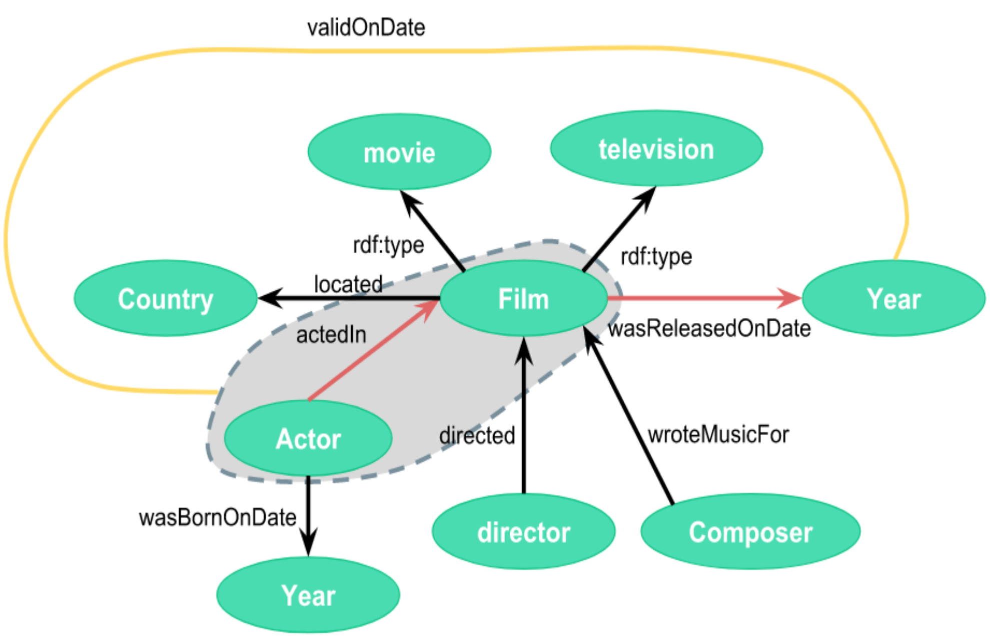
(A) TEMPORAL META-FACT GENERATION - PROPOSED APPROACH

[Malaverri et al. 2020]

Problem

- How can we capture fact validity time using only the facts and the structure of the KGs?

Approach – Seed meta-fact generation (step 1)



Graph Patterns

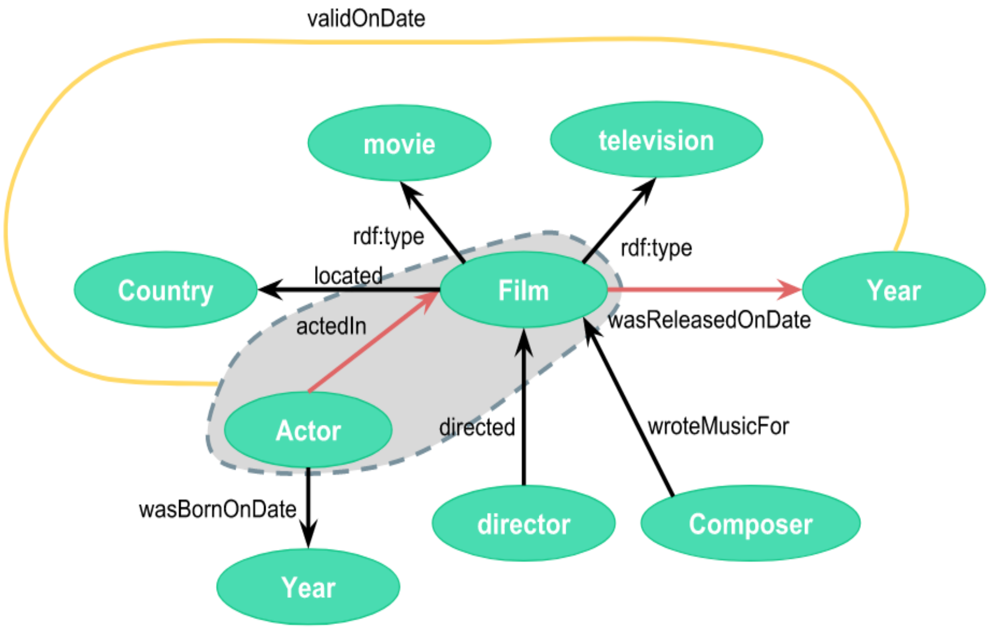
(A) TEMPORAL META-FACT GENERATION - PROPOSED APPROACH

[Malaverri et al. 2020]

Problem

- How can we capture fact validity time using only the facts and the structure of the KGs?

Approach – Seed meta-fact generation (step 1)



$$Actor(X) \wedge Film(y) \wedge actedIn(X, Y) \wedge$$
$$wasCreatedOnDate(X, d)$$
$$\Rightarrow inDateTime(actedIn(X, Y), d)$$

Graph Patterns

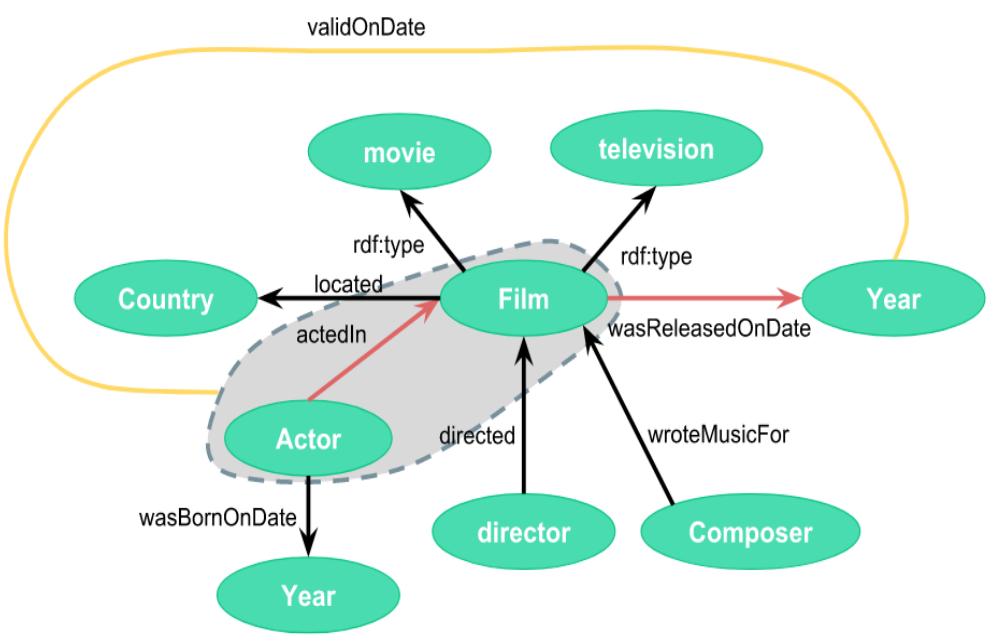
(A) TEMPORAL META-FACT GENERATION - PROPOSED APPROACH

[Malaverri et al. 2020]

Problem

- How can we capture fact validity time using only the facts and the structure of the KGs?

Approach – Seed meta-fact generation (step 1)



$$Actor(X) \wedge Film(y) \wedge actedIn(X, Y) \wedge$$
$$wasCreatedOnDate(X, d)$$
$$\Rightarrow inDateTime(actedIn(X, Y), d)$$

Temporal meta-fact

Graph Patterns

[Malaverri et al. 2020]

1. SEED META-FACT GENERATION

Algorithm: Query generation

- For each knowledge pattern, in the form of:

$$Actor(X) \wedge Film(y) \wedge actedIn(X, Y) \wedge wasCreatedOnDate(X, d)$$
$$\Rightarrow inDateTime(actedIn(X, Y), d)$$

1. SEED META-FACT GENERATION

Algorithm: Query generation

- For each knowledge pattern, in the form of:

$$Actor(X) \wedge Film(y) \wedge actedIn(X, Y) \wedge wasCreatedOnDate(X, d) \Rightarrow inDateTime(actedIn(X, Y), d)$$

- Generate a SPARQL query (graph pattern = rule premise)

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX yago: <http://yago-knowledge.org/resource/>
SELECT distinct ?a ?f ?d WHERE {
  ?a yago:actedIn ?f.
  ?f yago:wasCreatedOnDate ?d.
}
```

[Malaverri et al. 2020]

1. SEED META-FACT GENERATION

Algorithm: Meta-fact generation

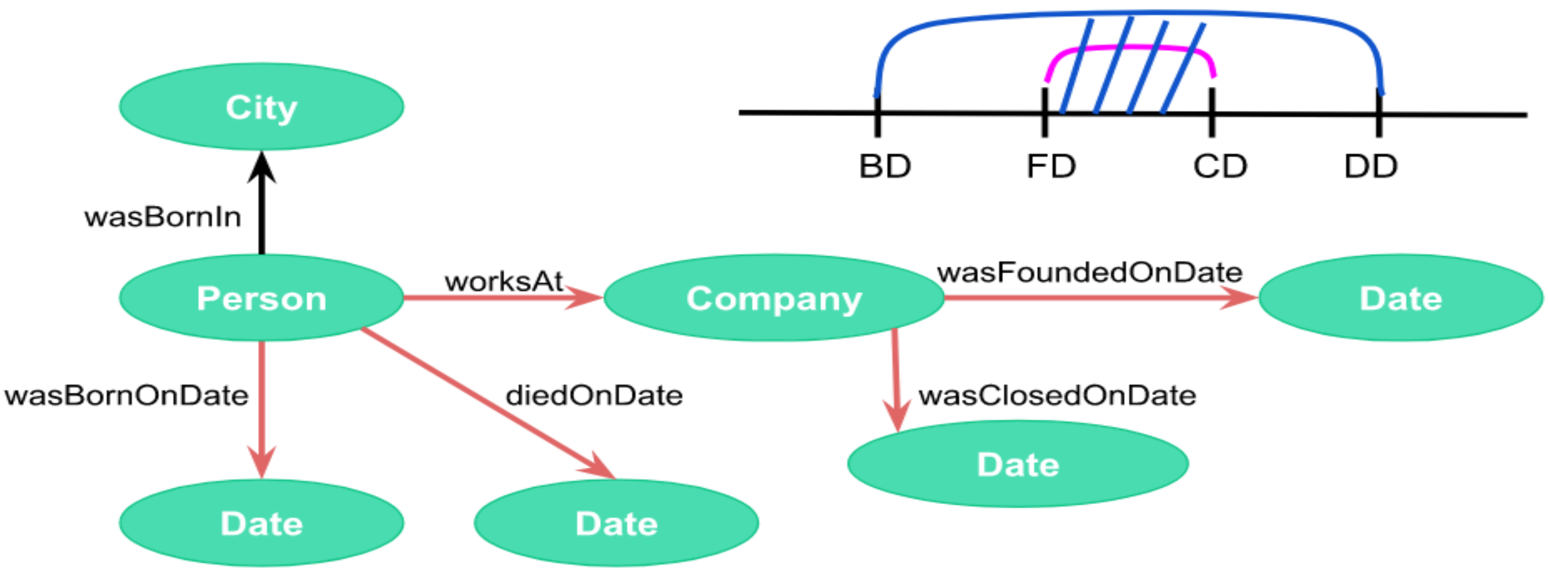
- On the set of facts obtained from SPARQL queries, generate the set of temporal meta-facts

[Malaveri et al. 2020]

1. SEED META-FACT GENERATION

Algorithm: Meta-fact generation

- On the set of facts obtained from SPARQL queries, generate the set of temporal meta-facts
- **For interval graph patterns**: compute an intersection between the temporal intervals associated to the subject and the object of a fact.

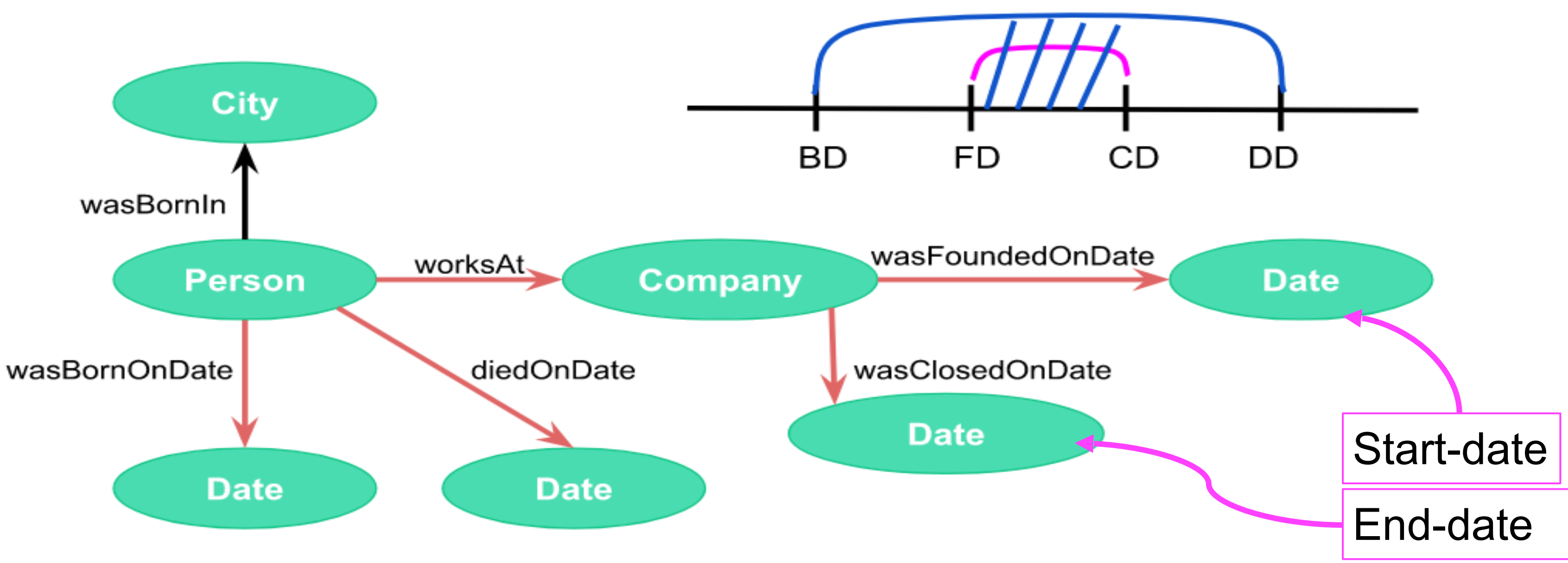


[Malaverri et al. 2020]

1. SEED META-FACT GENERATION

Algorithm: Meta-fact generation

- On the set of facts obtained from SPARQL queries, generate the set of temporal meta-facts
- **For interval graph patterns**: compute an intersection between the temporal intervals associated to the subject and the object of a fact.



TEMPORAL META-FACT EXPANSION

[Malaverri et al. 2020]

2. TEMPORAL META-FACT EXPANSION

- Use a set of Horn rules generated by a rule mining tool (e.g. AMIE) to propagate existing temporal meta-facts

$$\text{hasChild}(x, c) \land \text{isMarriedTo}(x, y) \implies \text{hasChild}(y, c)$$

[Malaveri et al. 2020]

2. TEMPORAL META-FACT EXPANSION

- Use a set of Horn rules generated by a rule mining tool (e.g. AMIE) to propagate existing temporal meta-facts

$$\text{hasChild}(x, c) \wedge \text{isMarriedTo}(x, y) \implies \text{hasChild}(y, c)$$

- **Step 1: Rule selection and instantiation**
  - Keep only rule with **confidence > threshold theta**
  - Select those premises contain a predicate in the set of seed meta-facts
  - Rule instantiation on a set of facts and seed meta-facts (generated or already existing in the KG)

[Malaverri et al. 2020]

2. TEMPORAL META-FACT EXPANSION

- Use a set of Horn rules generated by a rule mining tool (e.g. AMIE) to propagate existing temporal meta-facts

$$\text{hasChild}(x, c) \wedge \text{isMarriedTo}(x, y) \implies \text{hasChild}(y, c)$$

- **Step 2: Temporal meta-fact propagation**
  - Case 1: only one meta-fact in the premise, then propagate the date attached to the conclusion
  - Case 2: if several meta-facts in the premise, then apply temporal combination constraints

2. TEMPORAL COMBINATION CONSTRAINTS

[Malaverri et al. 2020]

- Based on Allen's calculus

	Time relations	Inferred time
1	during(TS, TI)	TI
2	before(TI_1, TI_2)	[start(TI_1) .. end(TI_2)]
3	during(TI_1, TI_2)	TI_2
4	overlaps(TI_1, TI_2)	[min(start(TI_1), start(TI_2)) .. max(end(TI_1), end(TI_2))]
5	meet(TI_1, TI_2)	[min(start(TI_1), start(TI_2)) .. max(end(TI_1), end(TI_2))]
6	before(TS_1, TS_2) and before(TS_2, TI)	[TS_1 .. end(TI)]
7	before(TI, TS_1) and before(TS_1, TS_2)	[start(TI) .. TS_2]
8	during(TS_1, TI) and before(TI, TS_2)	[start(TI) .. TS_2]
9	before(TI_1, TI_2) and before(TI_2, TI_3)	[start(TI_1) .. end(TI_3)]
10	during(TI_1, TI_2) and before(TI_2, TI_3)	[start(TI_1) .. end(TI_3)]
11	overlaps(TI_1, TI_2) and before(TI_2, TI_3)	[min(start(TI_1), start(TI_2)) .. end(TI_3)]
12	overlaps(TI_1, TI_2) and during(TI_2, TI_3)	TI_3

EXPERIMENTS

1. SEED META-FACT GENERATION: EXPERIMENTS ON YAGO

- **Yago3**: contains information extracted from Wikipedia infoboxes, WordNet, and GeoNames,
- **> 10 million** entities (persons, cities, organizations),
- **> 120 million** facts about these entities
- Attaches temporal and spatial dimensions to some facts and entities.

Predicates
actedIn
wroteMusicFor
created
participatedIn
graduatedFrom
isMarriedTo
hasAcademicAdvisor
isLeaderOf
playsFor
isAffiliatedTo
worksAt
hasChild
directed
edited

Time-sensitive predicates

1. SEED META-FACT GENERATION: EXPERIMENTS ON YAGO

Predicates	#YAGO MF	# Generated MF	
actedIn	36	305247	x 8479
wroteMusicFor	165	59130	
created	1943	444853	x 228
participatedIn	419	2655	x 6
graduatedFrom	3561	142137	
isMarriedTo	10168	53039	
hasAcademicAdvisor	95	9548	x 100
isLeaderOf	187	18471	
playsFor	34000	783254	
isAffiliatedTo	5290	783254	
worksAt	408	27611	x 67
hasChild	–	38236	
directed	–	90961	
edited	–	20008	

Quantitative results of seed meta-fact generation

2. TEMPORAL META-FACT EXPANSION: EXPERIMENTS

- DBpedia as a KG to be enriched with temporal meta-facts
- As seed meta-facts:
  - Yago meta-facts: **datasets A**
  - Seed meta-facts: **dataset B**
- Use of AMIE horn rules
- Wikidata as a **ground truth** (qualifiers)
  - Use of Wikidata endpoint *

Predicates	# of Records
hasChild	100000
directed	18726
created	100000
isMarriedTo	32275
isAffiliatedTo*	88087
worksAt	90000

* https://query.wikidata.org/

2. TEMPORAL META-FACT EXPANSION: EXPERIMENTS

- Quantitive results on **dataset A (Yago meta-facts)**

Predicates	# Original Meta-facts	# Meta-facts obtained	Rules applied	PCA Confidence	Head Coverage
hasChild	0	2156	?e<hasChild>?b,?e<isMarriedTo>?a =>?a<hasChild>?b	0.539998107	0.380672718
			?f<hasChild>?b?,a<isMarriedTo>?f =>?a<hasChild>?b	0.540100369	0.380672718
directed	0	1932	?a<created>?b=>?a<directed>?b	0.301768006	0.233104873
isMarriedTo	10168	8290	?b<isMarriedTo>?a=>?a<isMarriedTo>?b	0.999915811	0.999831636
isAffiliatedTo	5290	34000	?a<playsFor>?b=>?a<isAffiliatedTo>?b	0.999991338	0.824862929
playsFor	34000	5290	?a<isAffiliatedTo>?b=>?a<playsFor>?b	0.967279918	0.999991338
worksAt	408	177	?e<graduatedFrom>?b, ?e<hasAcademicAdvisor>?a =>?a<worksAt>?b	0.400990099	0.056577416

2. TEMPORAL META-FACT EXPANSION: EXPERIMENTS

- Quantitive results on **dataset B (Seed meta-facts, step1)**

Predicates	# Original Meta-facts	# Meta-facts obtained	Rules applied	PCA Confidence	Head Coverage
hasChild	38236	13231	?e<hasChild>?b,?e<isMarriedTo>?a =>?a<hasChild>?b	0.539998107	0.380672718
			?f<hasChild>?b?,a<isMarriedTo>?f =>?a<hasChild>?b	0.540100369	0.380672718
directed	90961	426187	?a<created>?b=>?a<directed>?b	0.301768006	0.233104873
created	444853	72620	?a<directed>?b=>?a<created>?b	0.392707051	0.035430598
isMarriedTo	53039	62416	?b<isMarriedTo>?a=>?a<isMarriedTo>?b	0.999915811	0.999831636
			?a<hasChild>?f,?b<hasChild>?f =>?a<isMarriedTo>?b	0.424990052	0.134859837
isAffiliatedTo	1204540	777039	?a<playsFor>?b=>?a<isAffiliatedTo>?b	0.999991338	0.824862929
worksAt	27611	5520	?e<graduatedFrom>?b, ?e<hasAcademicAdvisor>?a =>?a<worksAt>?b	0.400990099	0.056577416

2. TEMPORAL META-FACT EXPANSION: EXPERIMENTS

Qualitative results: Considered timestamp meta-facts and full date

		Dataset A			Dataset B		
		P	R	F	P	R	F
Confidence	0.1	0.89	0.89	0.89	0.67	0.67	0.67
	0.2	0.89	0.89	0.89	0.67	0.67	0.67
	0.3	0.89	0.89	0.89	0.67	0.67	0.67
	0.4	0.89	0.89	0.89	0.67	0.67	0.67
	0.5	0.89	0.89	0.89	0.67	0.67	0.67
	0.6	**0.89**	**0.89**	**0.89**	**0.67**	**0.66**	**0.66**
	0.7	**0.99**	**0.62**	**0.76**	**0.0008**	**0.0001**	**0.0002**
	0.8	0.99	0.62	0.76	0.0008	0.0001	0.0002
	0.9	0.99	0.62	0.76	0.0008	0.0001	0.0002

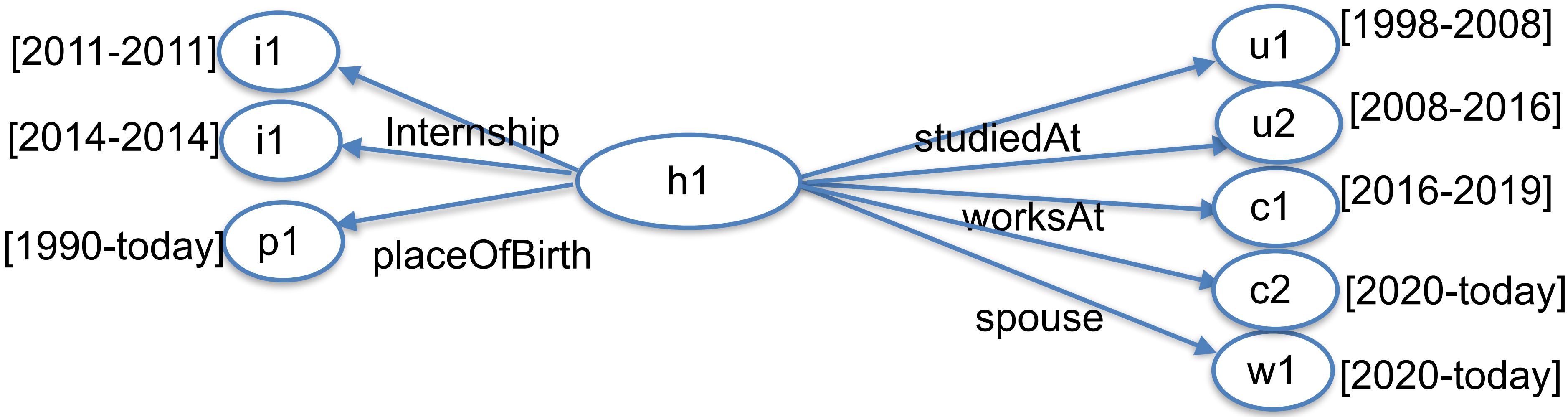
- **Dataset A** gets better results: better accuracy of seed meta-facts
- **Dataset B** reaches 0.67 of F-measure

CONCLUSION

- An approach for seed temporal meta-fact generation for time-sensitive predicates
- A rule-based approach for temporal meta-fact propagation
- A first quantitative and qualitative evaluation

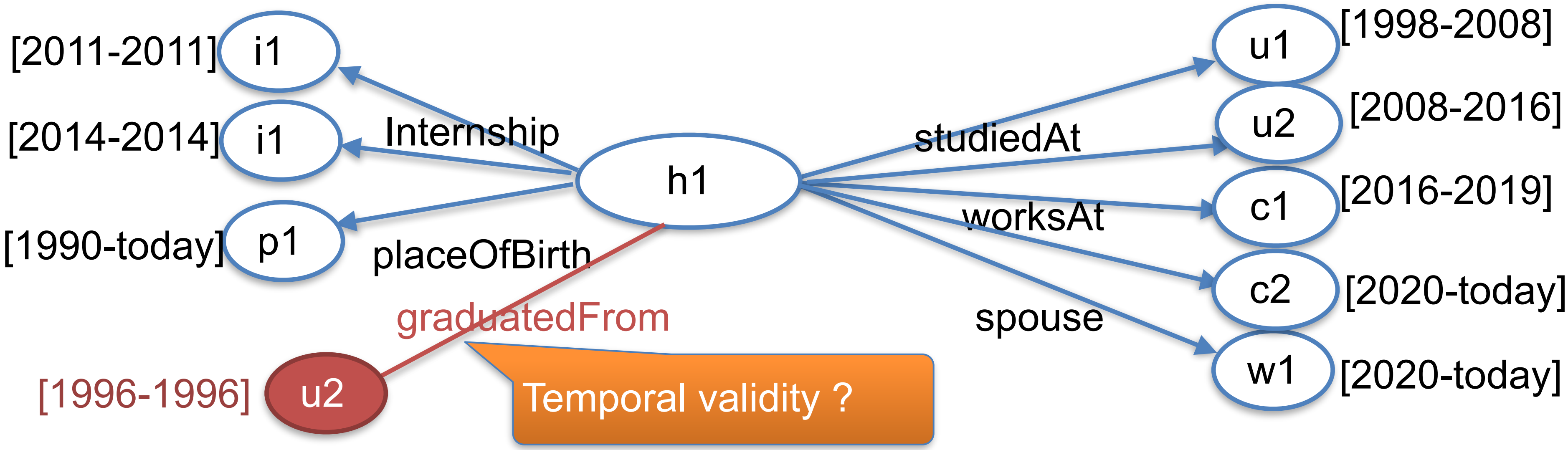
(B) HOW CAN WE ASSESS THE TEMPORAL VALIDITY OF FACTS

[Soulard et al. 2024]



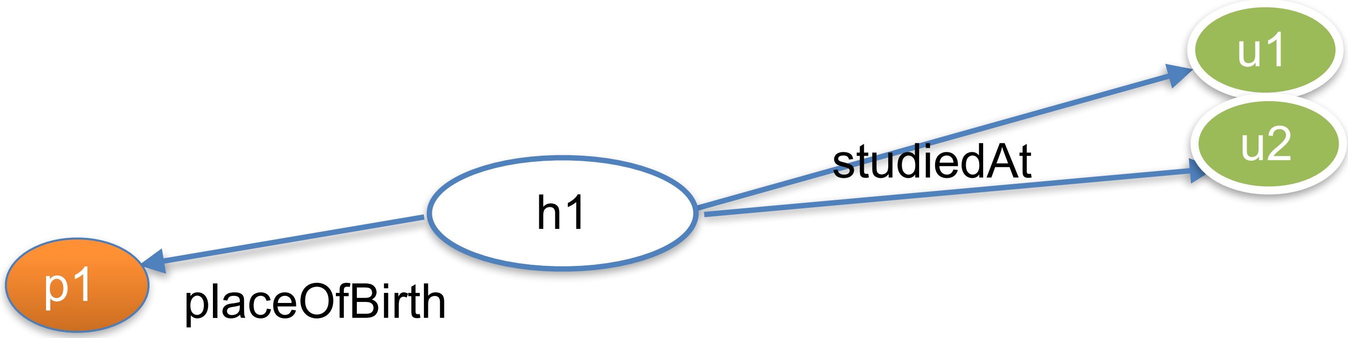
(B) HOW CAN WE ASSESS THE TEMPORAL VALIDITY OF FACTS

[Soulard et al. 2024]



(B) HOW CAN WE ASSESS THE TEMPORAL VALIDITY OF FACTS

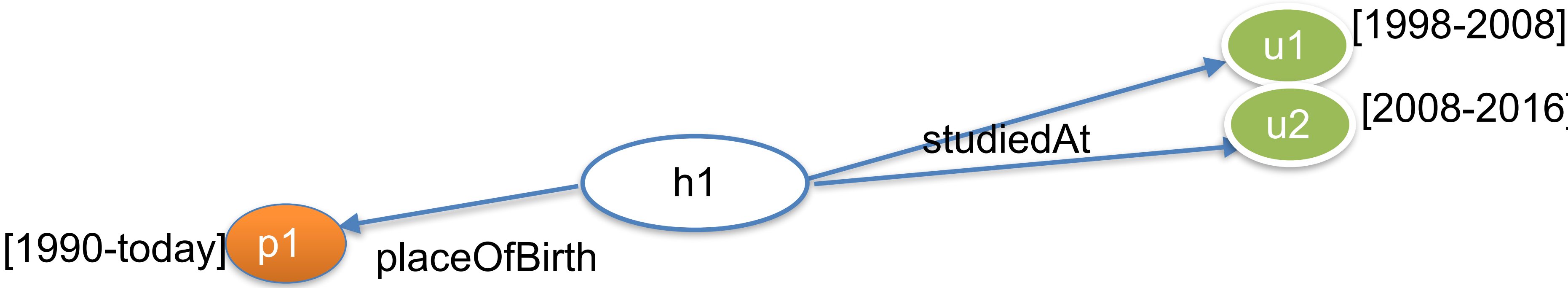
[Soulard et al. 2024]



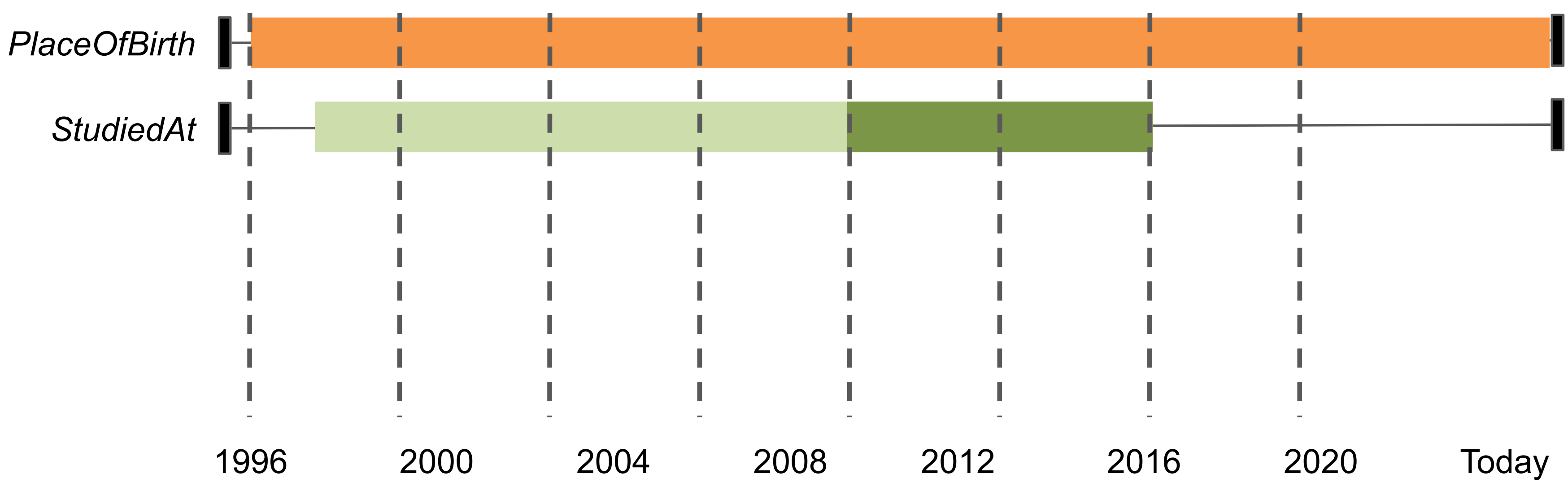
Entity, *h1* :

(B) HOW CAN WE ASSESS THE TEMPORAL VALIDITY OF FACTS

[Soulard et al. 2024]

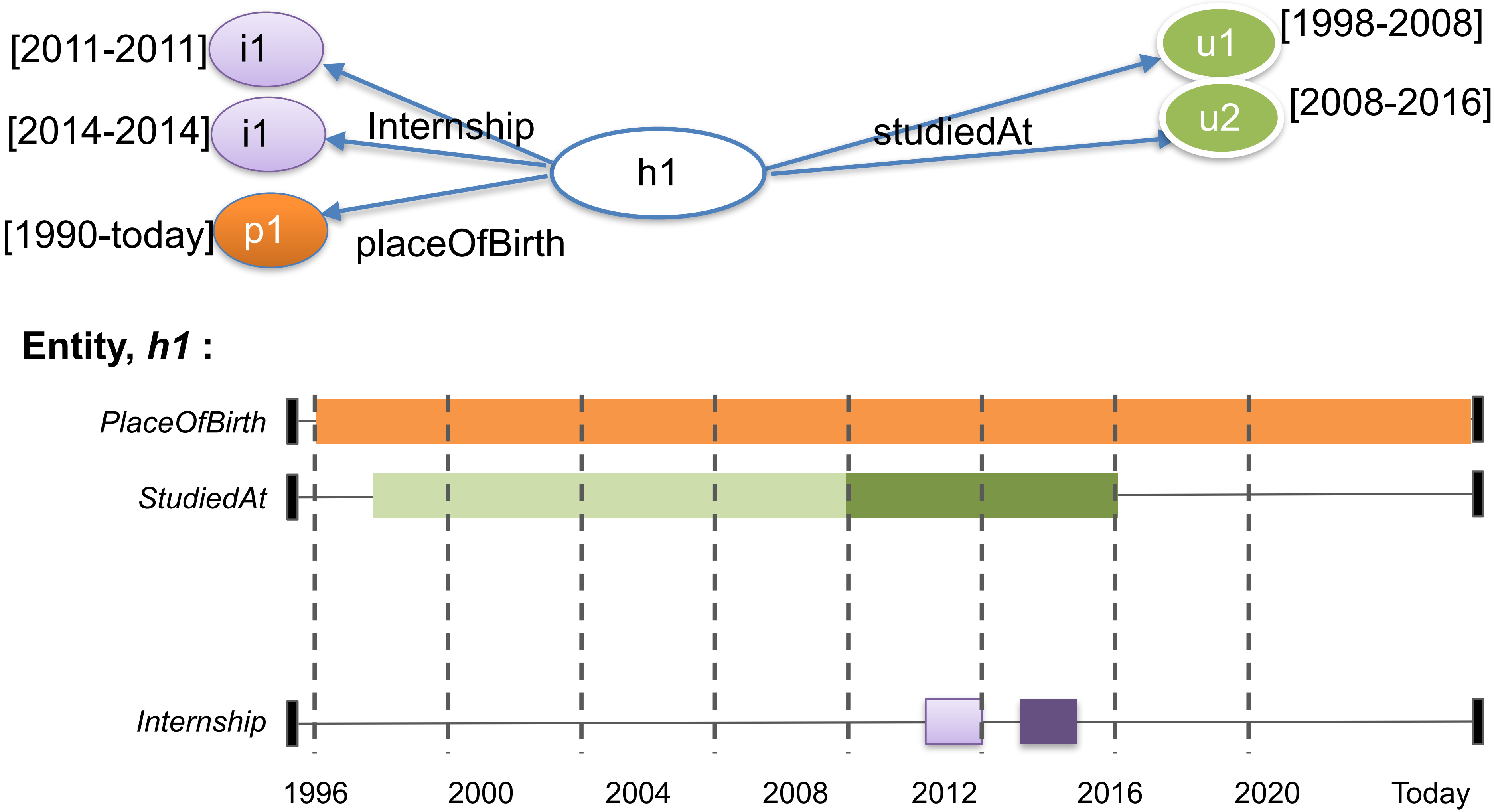


Entity, *h1* :



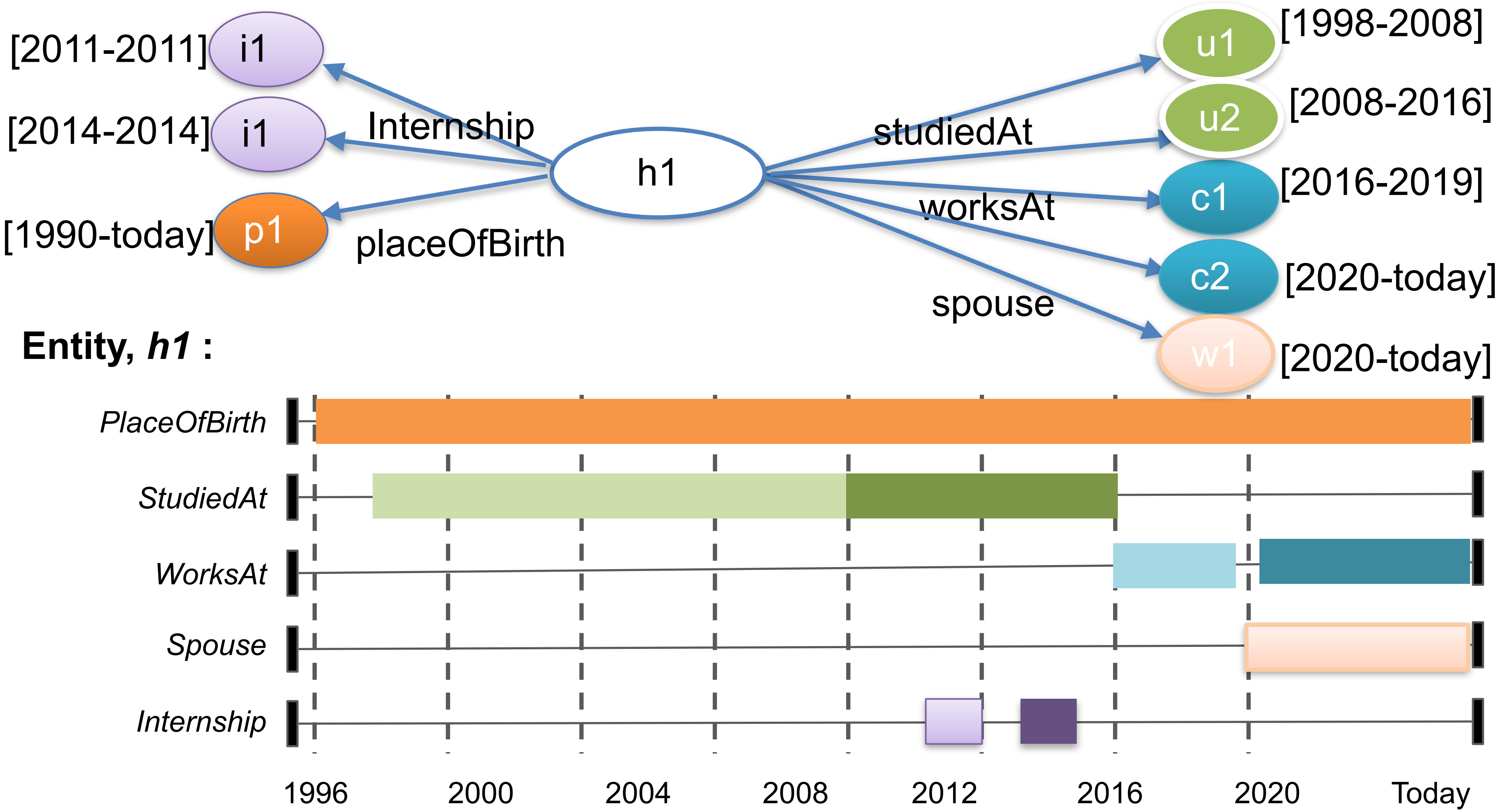
(B) HOW CAN WE ASSESS THE TEMPORAL VALIDITY OF FACTS

[Soulard et al. 2024]



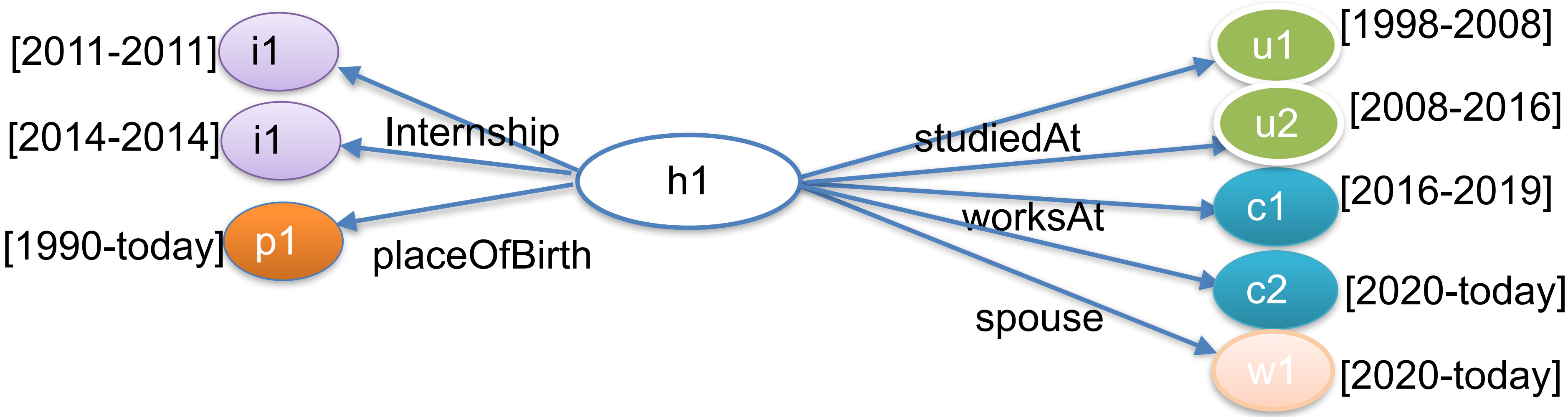
(B) HOW CAN WE ASSESS THE TEMPORAL VALIDITY OF FACTS

[Soulard et al. 2024]



(B) HOW CAN WE ASSESS THE TEMPORAL VALIDITY OF FACTS

[Soulard et al. 2024]



This graph represents several implicit temporal constraints

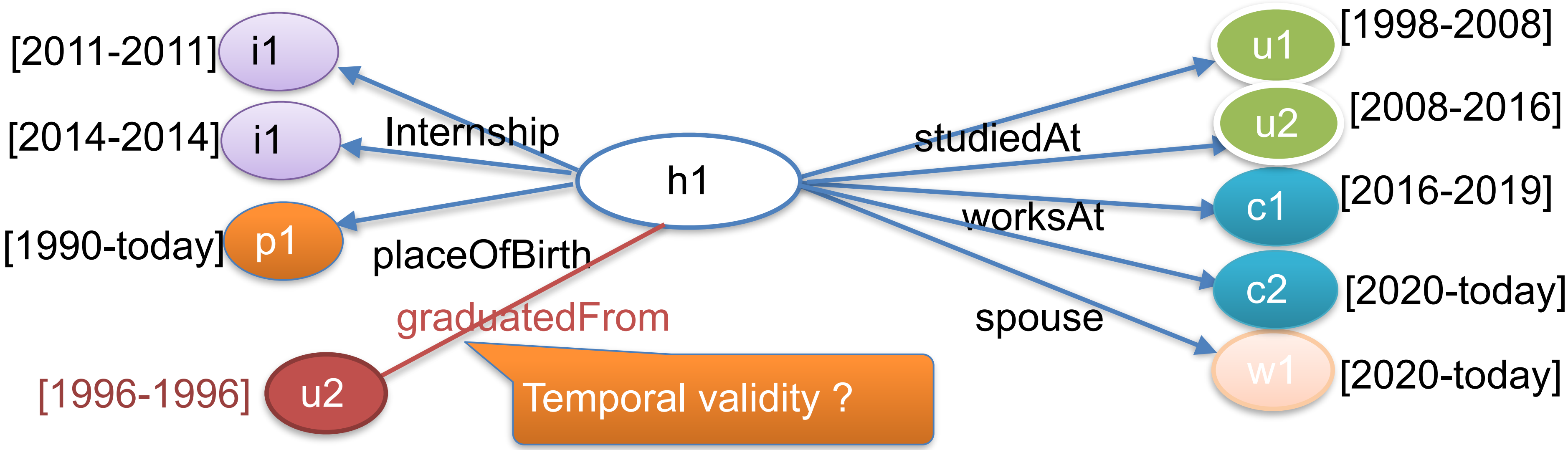
placeOfBirth [before] studiedIn

worksAt [after] Internship

internship [overlaps] studiedIn

(B) HOW CAN WE ASSESS THE TEMPORAL VALIDITY OF FACTS

[Soulard et al. 2024]



This graph represents several implicit temporal constraints

placeOfBirth [before] studiedIn

worksAt [after] Internship

internship [overlaps] studiedIn

(B) HOW CAN WE ASSESS THE TEMPORAL VALIDITY OF FACTS

[Soulard et al. 2024]

- **Problem:**
  - Let *f* be a fact in a temporal knowledge graph *KG* in the form of a quadruplet *(s, p, o, [sd, ed])* with *[sd, ed]* an intervalle of time representing the validity time of *f*.
  - *f* is valid in *KG* if the time intervalle is temporally consistant with respect to the temporal constraints fulfilled in *KG.*

(B) HOW CAN WE ASSESS THE TEMPORAL VALIDITY OF FACTS

[Soulard et al. 2024]

- **Problem:**
  - Let *f* be a fact in a temporal knowledge graph *KG* in the form of a quadruplet *(s, p, o, [sd, ed])* with *[sd, ed]* an intervalle of time representing the validity time of *f*.
  - *f* is valid in *KG* if the time intervalle is temporally consistant with respect to the temporal constraints fulfilled in *KG.*

Our approach

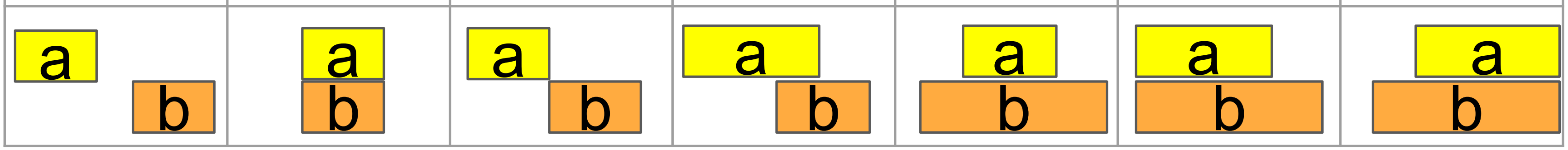
1. Temporal constraint discovery
2. Temporal validity checking of facts with respect to the discovered temporal constraints

B.1. TEMPORAL CONSTRAINT DISCOVERY

[Soulard et al. 2024]

- **Allen's Algebra temporal relations between simple intervals**

Before	Equals	Meets	Overlaps	During	Starts	Finishes

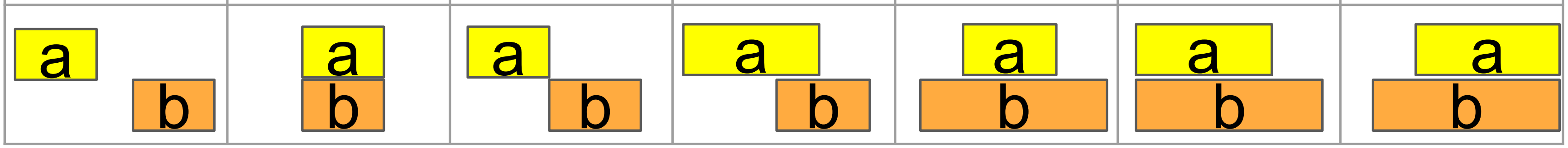


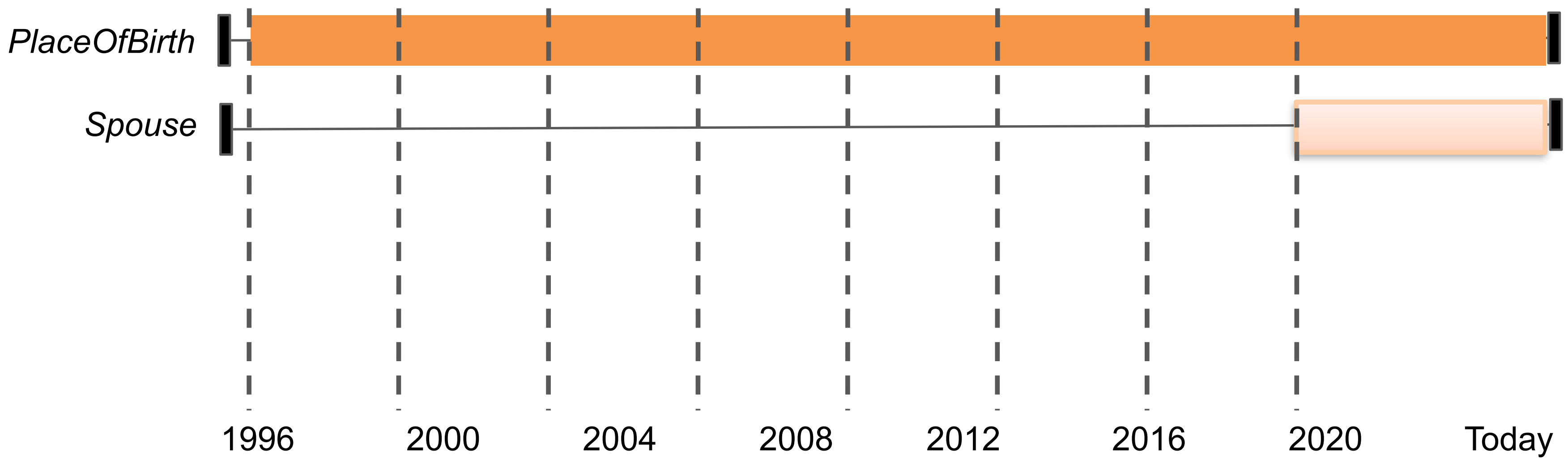
B.1. TEMPORAL CONSTRAINT DISCOVERY

[Soulard et al. 2024]

- **Allen's Algebra temporal relations between simple intervals**

Before	Equals	Meets	Overlaps	During	Starts	Finishes



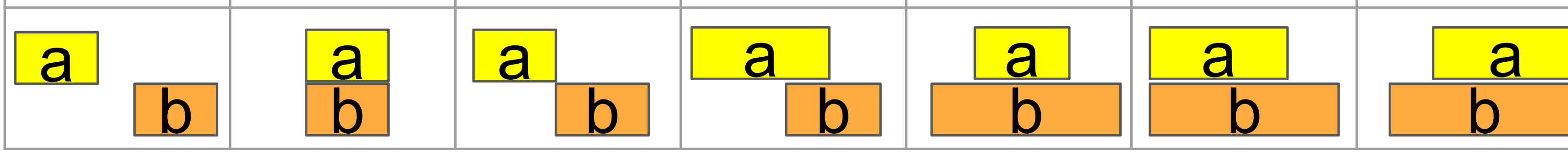


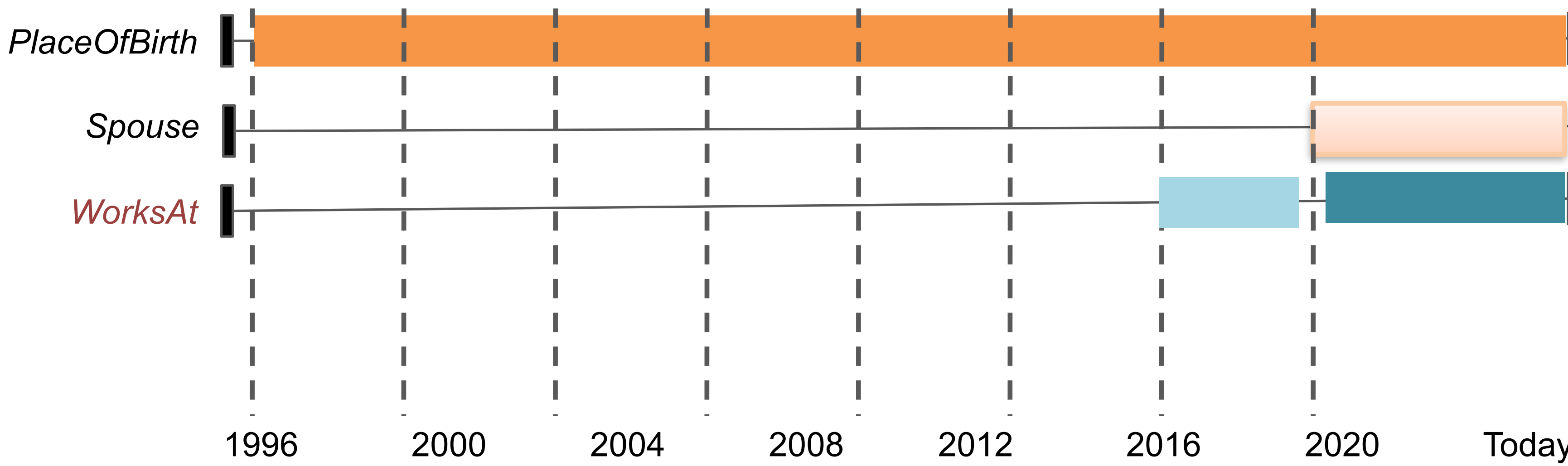
B.1. TEMPORAL CONSTRAINT DISCOVERY

[Soulard et al. 2024]

- **Allen's Algebra temporal relations between simple intervals but not for sequences of intervals**

Before	Equals	Meets	Overlaps	During	Starts	Finishes

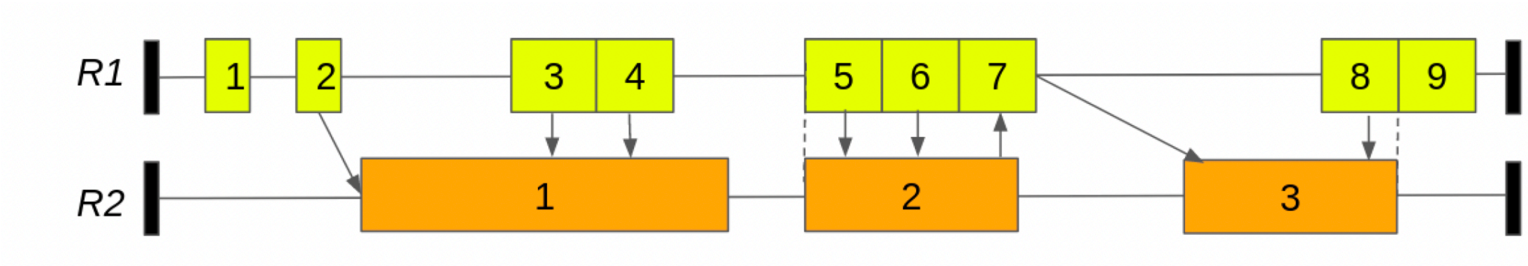




[Soulard et al. 2024]

B.1. TEMPORAL CONSTRAINT DISCOVERY - EXTENSION TO SEQUENCES

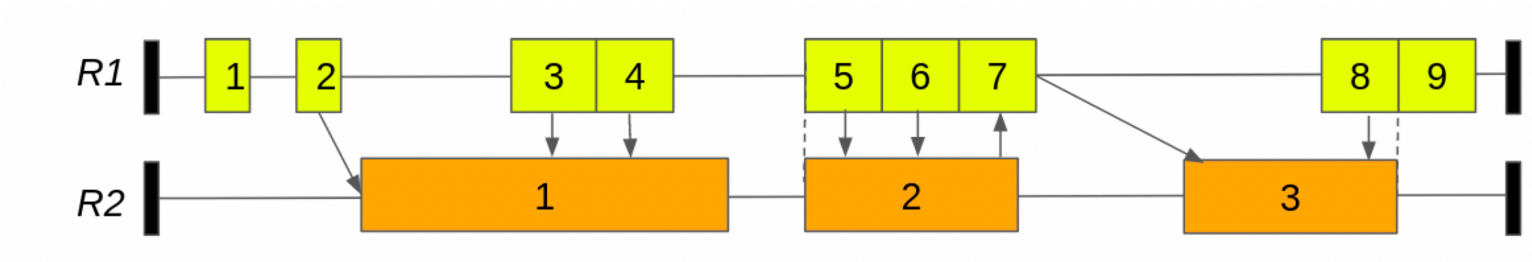
- New algorithm for sequence comparison that deals with intra and inter-comparisons in one passe



[Soulard et al. 2024]

B.1. TEMPORAL CONSTRAINT DISCOVERY - EXTENSION TO SEQUENCES

- New algorithm for sequence comparison that deals with intra and inter-comparisons in one passe



Relation	$o(S_1.I, S_1.I')$	$o(S_2.I, S_2.I')$
Meets	4	0

Intra-sequences comparisons $M_{\triangleleft}$

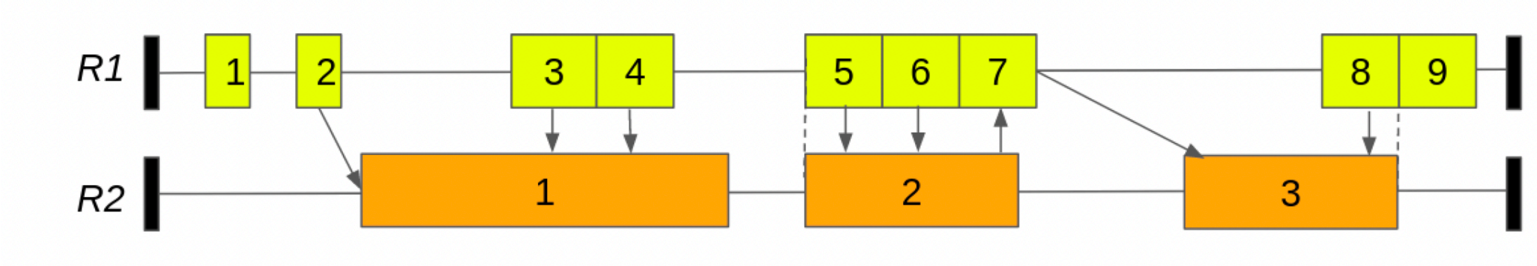
Relation	$o(S_1.I, S_2.I)$	$o(S_2.I, S_1.I)$
Before	2	0
Equals	0	0
Meets	0	0
Overlaps	0	1
During	3	0
Starts	1	0
Finishes	1	0

Inter-sequences comparisons $M_{\triangleright}$

[Soulard et al. 2024]

B.1. TEMPORAL CONSTRAINT DISCOVERY - EXTENSION TO SEQUENCES

- New algorithm for sequence comparison that deals with intra and inter-comparisons in one passe



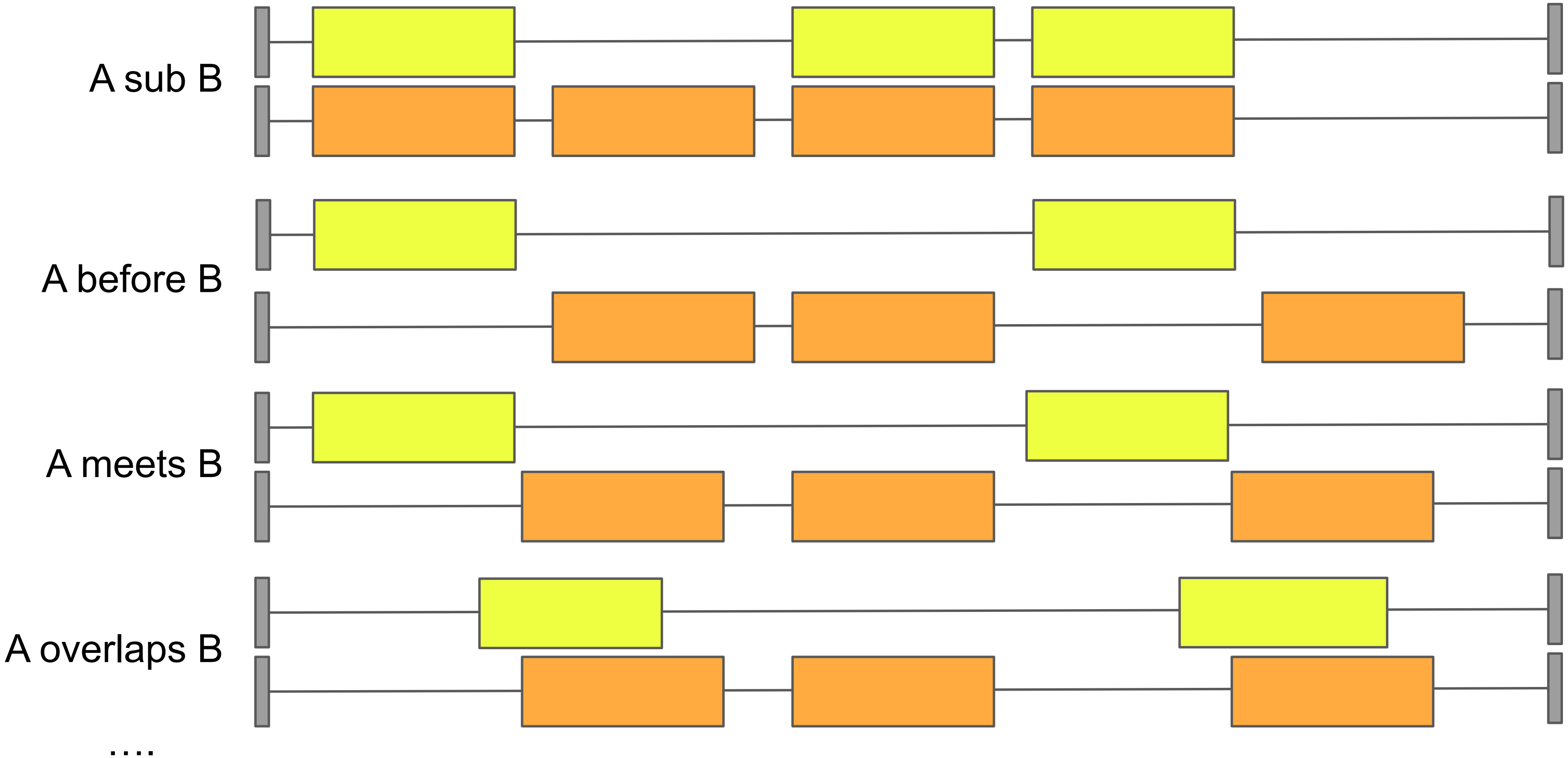
Relevant Inter-sequence comparisons

$$
\begin{aligned}
&(I \cap_t I' \neq \emptyset) \\
\vee\ &(I.s < I'.e) \\
&\quad \wedge(\nexists I'' \in S \setminus \{I\}, (I''.s \geq I.e \wedge I''.s \leq I'.s) \\
&\quad \wedge(\nexists I'' \in S' \setminus \{I'\}, (I''.e \leq I'.e \wedge I''.e \geq I.s) \\
\vee\ &(I.s > I'.e) \\
&\quad \wedge(\nexists I'' \in S \setminus \{I\}, (I''.e \geq I'.s \wedge I''.e \leq I.s) \\
&\quad \wedge(\nexists I'' \in S' \setminus \{I'\}, (I''.s \leq I.s \wedge I''.s \geq I'.e),
\end{aligned}
$$

[Soulard et al. 2024]

B.1. TEMPORAL CONSTRAINT DISCOVERY - EXTENSION TO SEQUENCES

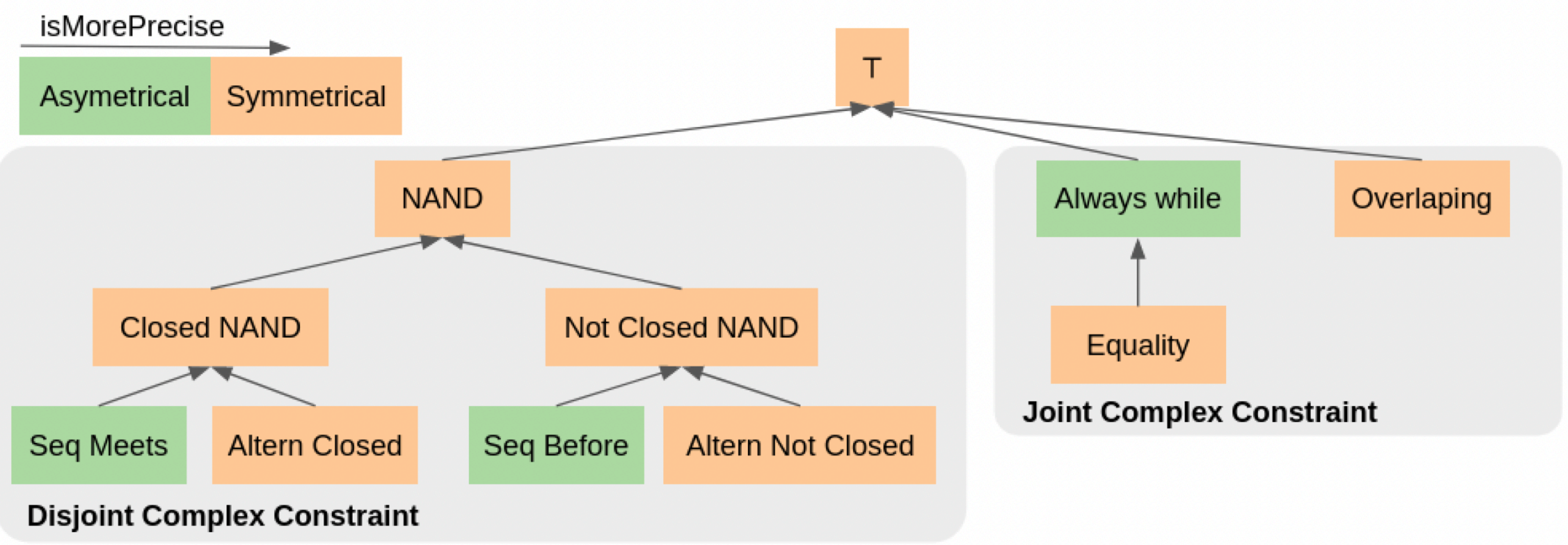
- We obtain simple Temporal Constraint, only composed of one type of relations (before, meets, overlaps, …)



[Soulard et al. 2024]

B.1. TEMPORAL CONSTRAINT DISCOVERY - EXTENSION TO SEQUENCES

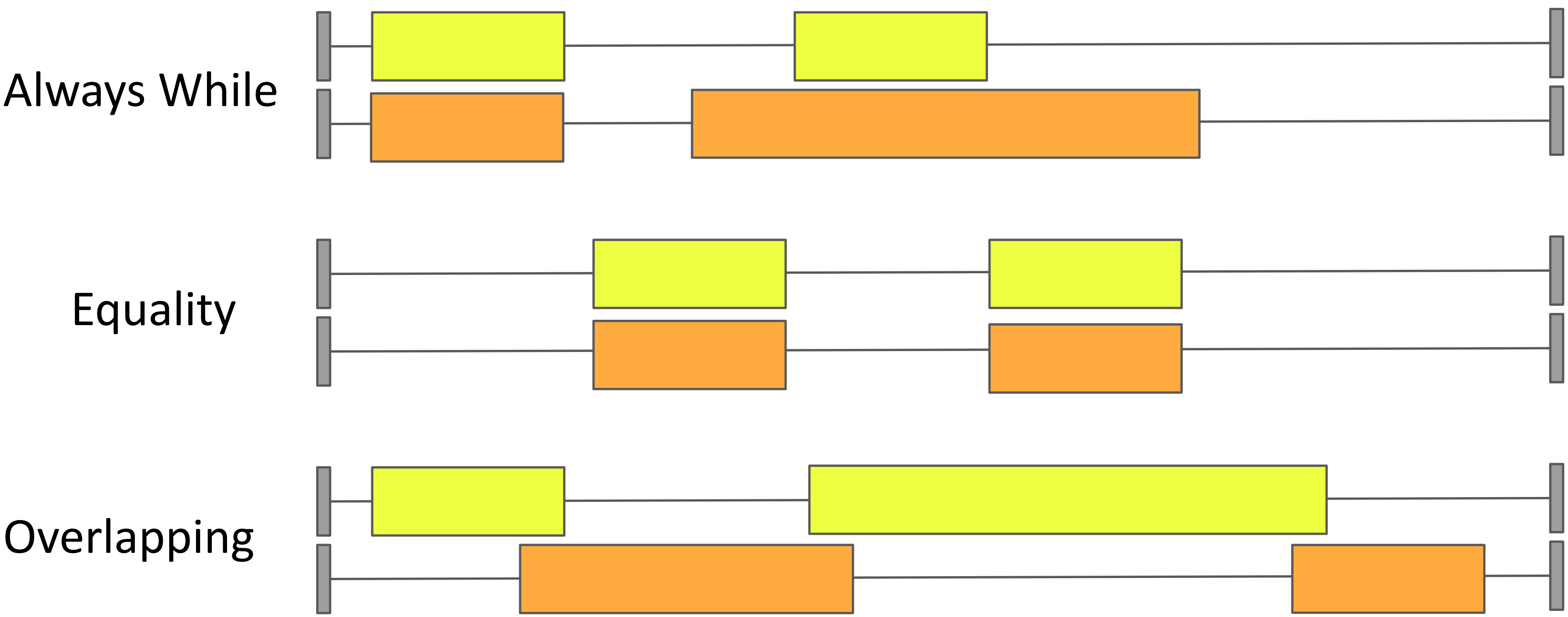
- We can then combine these simple constraints to represent complex ones composed of multiple types of relations or lack of.



[Soulard et al. 2024]

B.1. TEMPORAL CONSTRAINT DISCOVERY - EXTENSION TO SEQUENCES

- We can then combine these simple constraints to represent complex ones composed of multiple types of relations or lack of.



[Soulard et al. 2024]

B.1. TEMPORAL CONSTRAINT DISCOVERY - EXTENSION TO SEQUENCES

- We can then combine these simple constraints to represent complex ones composed of multiple types of relations or lack of.

[Soulard et al. 2024]

B.1. TEMPORAL CONSTRAINT DISCOVERY - GENERALIZATION TO CLASS LEVEL

For the entity ***h1***, we have discovered :
Seq.Internship **Always While** Seq.StudiedAt

But is the case for other (~all) entities of the Class ***Human*** ?

Generalisation Threshold

Is the constraint general enough to be used on a minimum of entities.

$$GeneRate(TC) = \frac{|E_{P,P'}|}{|E|}$$

Error Threshold

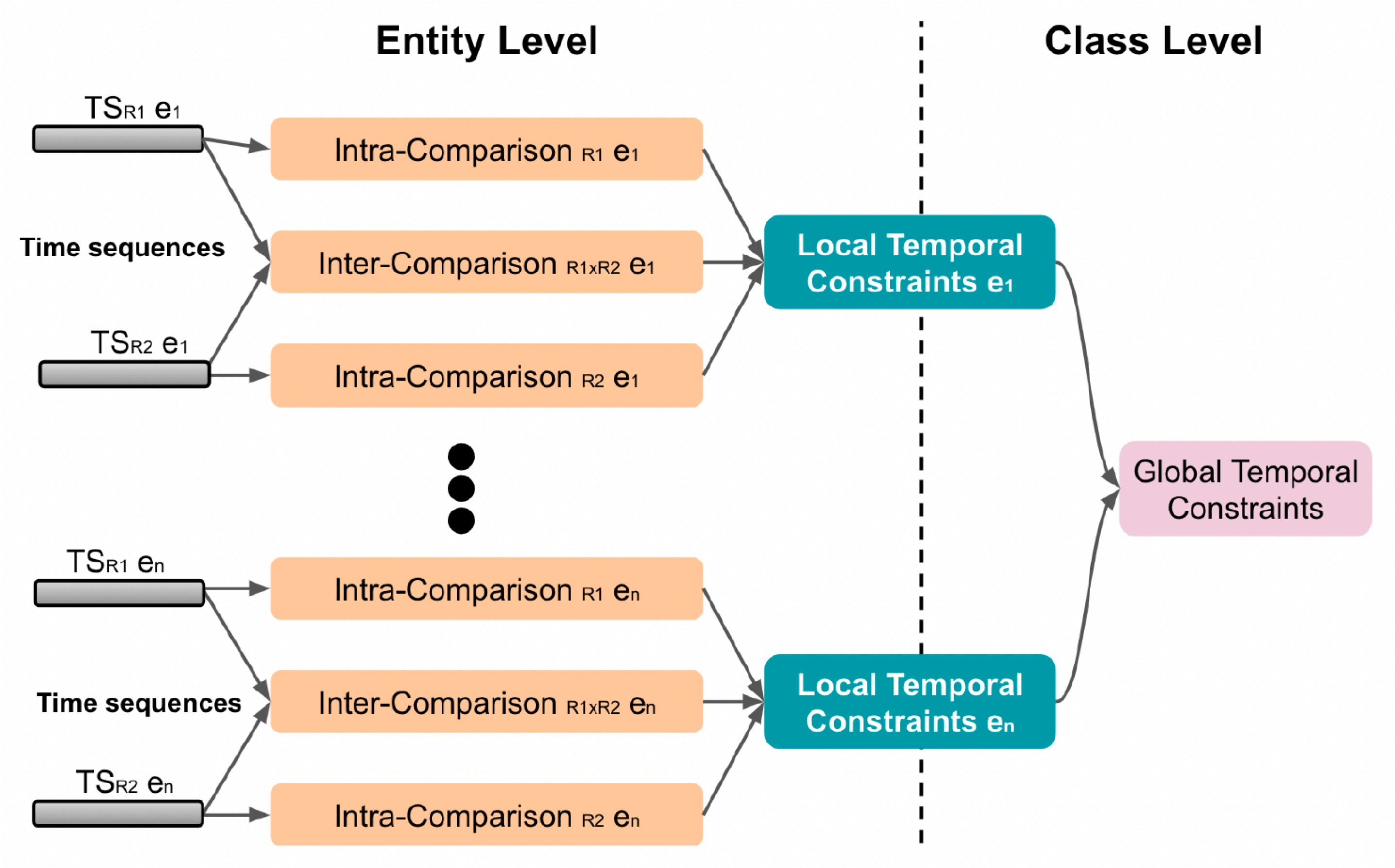
Is the constraint shared among the entities described by both relations.

$$ErrorRate(TC) = \frac{|X_{P,P'}|}{|E_{P,P'}|}$$

[Soulard et al. 2024]

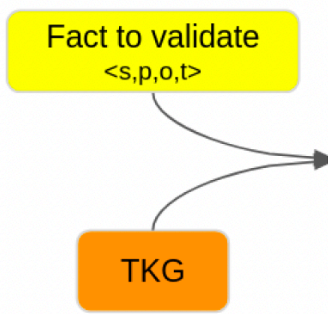
B.1. TEMPORAL CONSTRAINT DISCOVERY - EXTENSION TO SEQUENCES

Constraint Discovery Framework for two properties R1 and R2



[Soulard et al. 2024]

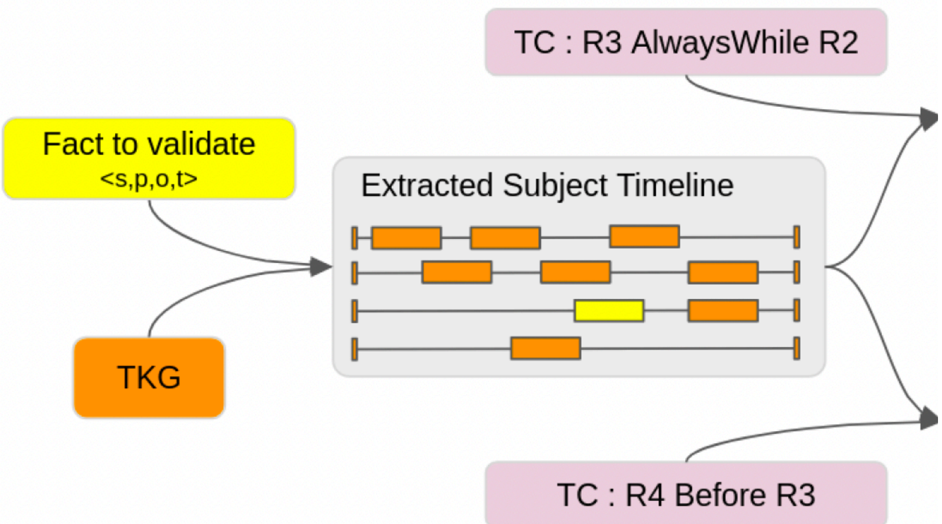
B.2. TEMPORAL FACT VALIDATION USING TEMPORAL CONSTRAINTS



- Constraint based Temporal Fact Validation framework -

[Soulard et al. 2024]

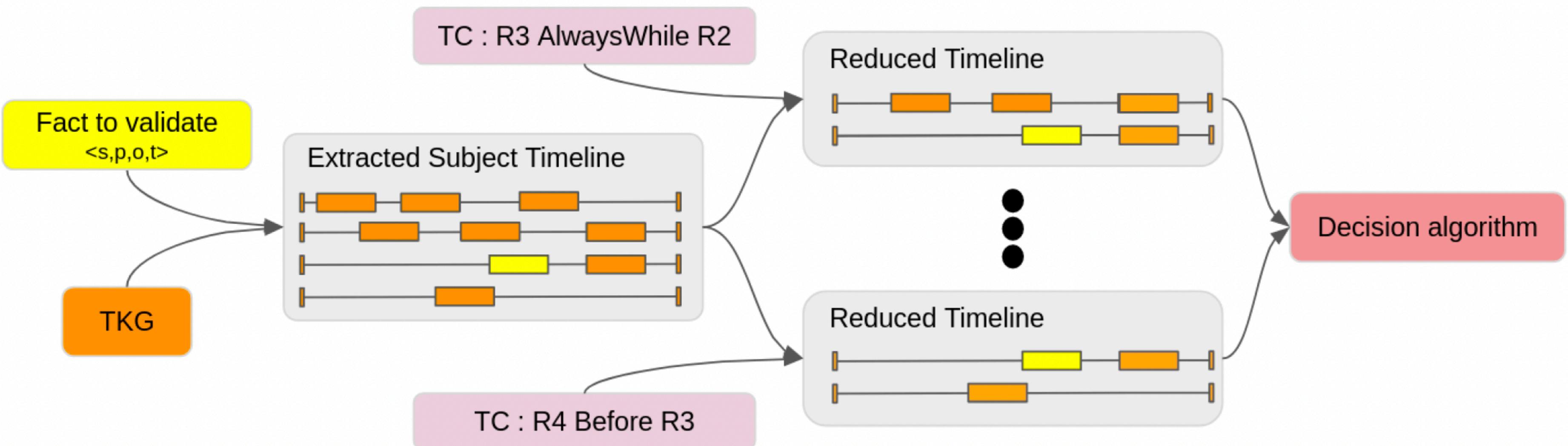
B.2. TEMPORAL FACT VALIDATION USING TEMPORAL CONSTRAINTS



- Constraint based Temporal Fact Validation framework -

[Soulard et al. 2024]

B.2. TEMPORAL FACT VALIDATION USING TEMPORAL CONSTRAINTS



- Constraint based Temporal Fact Validation framework -

[Soulard et al. 2024]

B.2. TEMPORAL FACT VALIDATION USING TEMPORAL CONSTRAINTS
- DECISION ALGORITHM

Symbolic approach - voting-based strategy : we check the temporal validity of a fact w.r.t all its relevant temporal constraints.

- Using a simple voting system w.r.t a given threshold

[Soulard et al. 2024]

B.2. TEMPORAL FACT VALIDATION USING TEMPORAL CONSTRAINTS
- DECISION ALGORITHM

Symbolic approach - voting-based strategy : we check the temporal validity of a fact w.r.t all its relevant temporal constraints.

- Using a simple voting system w.r.t a given threshold

Neuro-Symbolic combination strategy: we check the temporal validity of a fact w.r.t all available temporal constraints

- Then, for each temporal constraint tc, we associate a number to represent all possible behaviors (0, 1 and multiple values to represent the various possible cases)

- This results in a **matrix n$*$m,** where n is the size of the set of temporal constraints and m is the number of facts to validate, and a **ground truth vector** of dimension n that can be used to train and test the machine learning model (decision tree)

[Soulard et al. 2024]

B.2. TEMPORAL FACT VALIDATION USING TEMPORAL CONSTRAINTS

Datasets extracted from **Wikidata** (dec. 2023) representing all facts related to entities of different types

Class	# Entities	# Quadruplets
Country (Q6256)	205	183 249
Musical Group (Q215380)	55 507	131 476
Politician (Q82955)	658 445	2 085 232

[Soulard et al. 2024]

B.2. TEMPORAL FACT VALIDATION USING TEMPORAL CONSTRAINTS

Constraint discovery: Hyperparameters gen & err results for generalization step

Class	gen	err	Acc.	Cov.	R.T
Country	2	10	-	-	-
	2	5	**88.6**	16	8h 50m
	5	10	87.9	**17.2**	4h 30m
	5	5	88.5	14.4	**2h 54m**
Musical Group	2	10	**64.6**	**38.2**	6m 6s
	2	5	**64.6**	**38.2**	5m 54s
	5	10	64.2	37.5	**5m 51s**
	5	5	64.2	37.5	**5m 51s**
Politician	2	10	63.4	**51.9**	1h 35m
	2	5	62.1	49.9	**1h 32m**
	5	10	**63.5**	48.9	1h 34m
	5	5	62.1	46.9	**1h 32m**

[Soulard et al. 2024]

B.2. TEMPORAL FACT VALIDATION USING TEMPORAL CONSTRAINTS

Decision algorithms results

Class	Deci. Type	Acc.	Cov.	R.T	Size
Country	Symbolic	79.5	9.4	**2m 30s**	183K quads
	Neuro-Symb.	**80.4**	9.4	52m	
Musical Group	Symb.	64.0	37.5	**2m 50s**	131K quads
	Neuro-Symb.	**64.3**	37.5	5m 50s	
Politician	Symb.	61.6	44.0	**44m**	2M quads
	Neuro-Symb.	**62.3**	44.0	1h 30m	

[Soulard et al. 2024]

CONCLUSION

- **Good accuracy but limited coverage**
- **Limitations of the proposed approach :**
  - Case of no temporal constraints can be applied for a fact
  - If a relation is **temporally independant**
  - If multiple values are defined for the same relation at the **same time** (e.g. studying in several universities at the same time)

FUTURE WORK

- **Experiments** :
  - reduce the number of features that highly impact the performance of the neurosymbolic approach, by discarding temporal constraints that are less important.
  - evaluate whether the set of temporal constraints discovered in one graph can be transferred and used to validate or refute temporal facts on several graphs with high accuracy and coverage.
  - ➡ test the transferability of these temporal constraints on several other temporal KGs, such as YAGO

- **Methodology** : Use of knowledge graph embeddings by introducing temporal constraints in a knowledge graph embedding Loss.

REFERENCES

- *[Soulard et al. 2024] Thibaut Soulard, Joe Raad and Fatiha Saïs. Validation Temporelle Explicable de Faits par la Découverte de Contraintes Temporelles Complexes dans les Graphes de Connaissances. Ingénierie des connaissances (IC@PFIA 2024)*

- *[Malaverri et al.] Fatiha Saïs, Joana E. Gonzales Malaverri, Gianluca Quercini: MOMENT: temporal meta-fact generation and propagation in knowledge graphs. SAC 2020: 2039-2048*

- *[Gutierrez et al] Claudio Gutierrez, Carlos Hurtado and Alejandro Vaisman. Introducing Time into RDF. IEEE Transactions on Knowledge and Data Engineering. Vol. 19, Number. 2. 2007*

- *[Zhang et al] Fu Zhang, Ke Wang, Zhiyin Li and Jingwei Cheng. Temporal Data Representation and Querying Based on RDF. IEEE Access Journal*

- *[Foffart et al] Johannes Hoffart, Fabian M. Suchanek, Klaus Berberich, Gerhard Weikum. YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia. Artificial Intelligence, 2013*

- *[RDF*] https://w3c.github.io/rdf-star/cg-spec/editors_draft.html*

- *[Barbieri et al] D.F Barbieri, D. Braga, S. Ceri, E. Della Valle and M. Grossniklaus. C-SPARQL: A Continuous Query Language for RDF Data Streams. International Journal of Semantic Computing. Vol 04, Numbrer 1, 2010.*

Explainable Temporal Fact Validation Through Temporal Constraints Discovery In Knowledge Graphs

Thanks

FATIHA SAÏS

Joint Work With: JOANA E. GONZALES MALAVERRI, THIBAUT SOULARD AND JOE RAAD

LISN, CNRS & Université Paris Saclay

Séminaire SESAME - INRAE - 10/06/2024

cnrs dépasser les frontières

LISN LABORATOIRE INTERDISCIPLINAIRE DES SCIENCES DU NUMÉRIQUE

université PARIS-SACLAY