Fluffy '18
DTAI
Declarative Languages & Artificial Intelligence

Data Wrangling Automation

Lidia Contreras Ochando
liconoc@upv.es
@liconoc

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

DSIC
DEPARTAMENT DE SISTEMES INFORMÀTICS I COMPUTACIÓ

ELP GROUP

dmip

Approaches

1

Domain-Specific Induction

Joint work with José Hernández Orallo, Cèsar Ferri, Fernando Martínez Plumed, María José Ramírez Quintana and Susumu Katayama

Submitted to CIKM ‘18

2

Dynamic Background Knowledge

Joint work with José Hernández Orallo, Cèsar Ferri, Fernando Martínez Plumed, María José Ramírez Quintana and Susumu Katayama

Submitted to NIPS ‘18

3

Adaptive Domain Detection

Joint work with Gust Verbruggen, Luc De Raedt, José Hernández Orallo and Cèsar Ferri

Work in progress

Motivation

Example

Problem: Automate the transformation of data presented in different formats using few examples

Input data **Expected output**

Dates

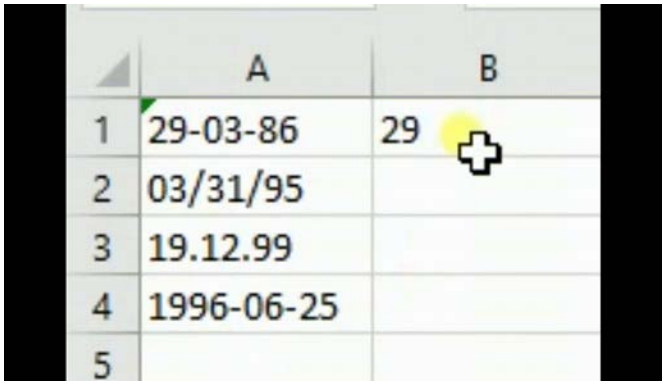
29-03-86 → 29

03/31/95 → 31

19.12.99 → 19

1996-06-25 → 25

Extract the day



	A	B
1	29-03-86	29
2	03/31/95	
3	19.12.99	
4	1996-06-25	
5		

FlashFill

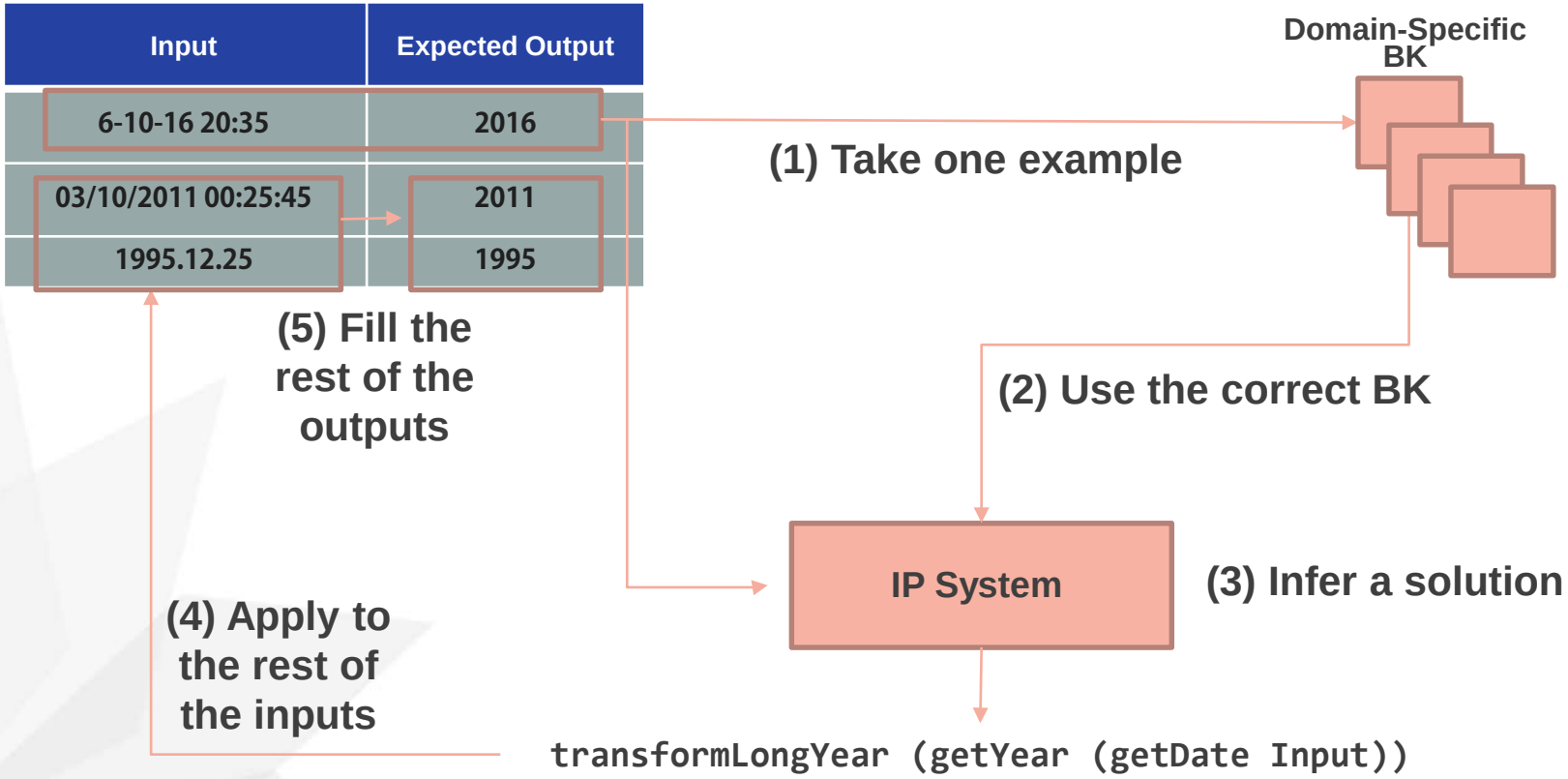
Specific functions to deal with dates (Background Knowledge)

Domain Specific Induction

Joint work with José Hernández Orallo, Cèsar Ferri, Fernando Martínez Plumed, María José Ramírez Quintana and Susumu Katayama
Submitted to CIKM '18

1. Domain-Specific Induction

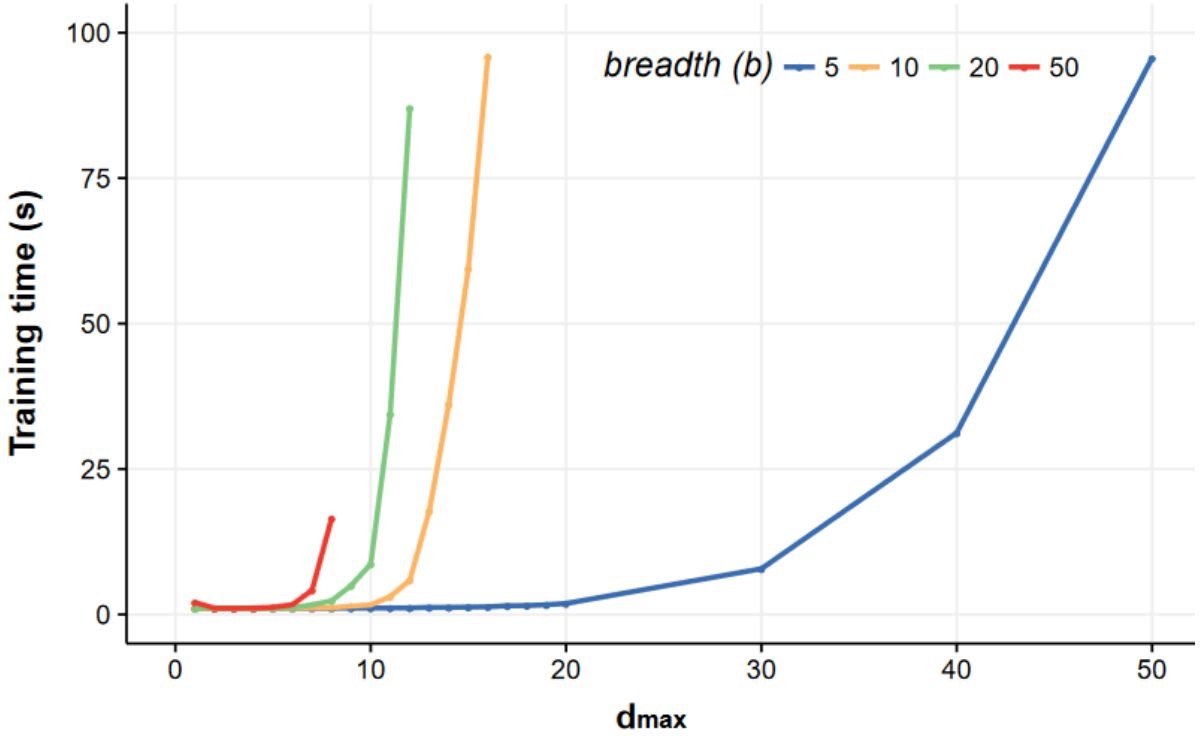
System functionality



1. Domain-Specific Induction

Limitations

- We need the user to select the domain.
- We can have a combinatorial explosion with big BKs.



Time needed for training depending on the maximum number of primitives that are allowed in any synthesised function (dmax) and the number of primitives in the BK (b).

2

Dynamic Background Knowledge

Joint work with José Hernández Orallo, Cèsar Ferri, Fernando Martínez Plumed, María José Ramírez Quintana and Susumu Katayama
Submitted to NIPS ‘18

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

宮崎大学 University of Miyazaki

dmip

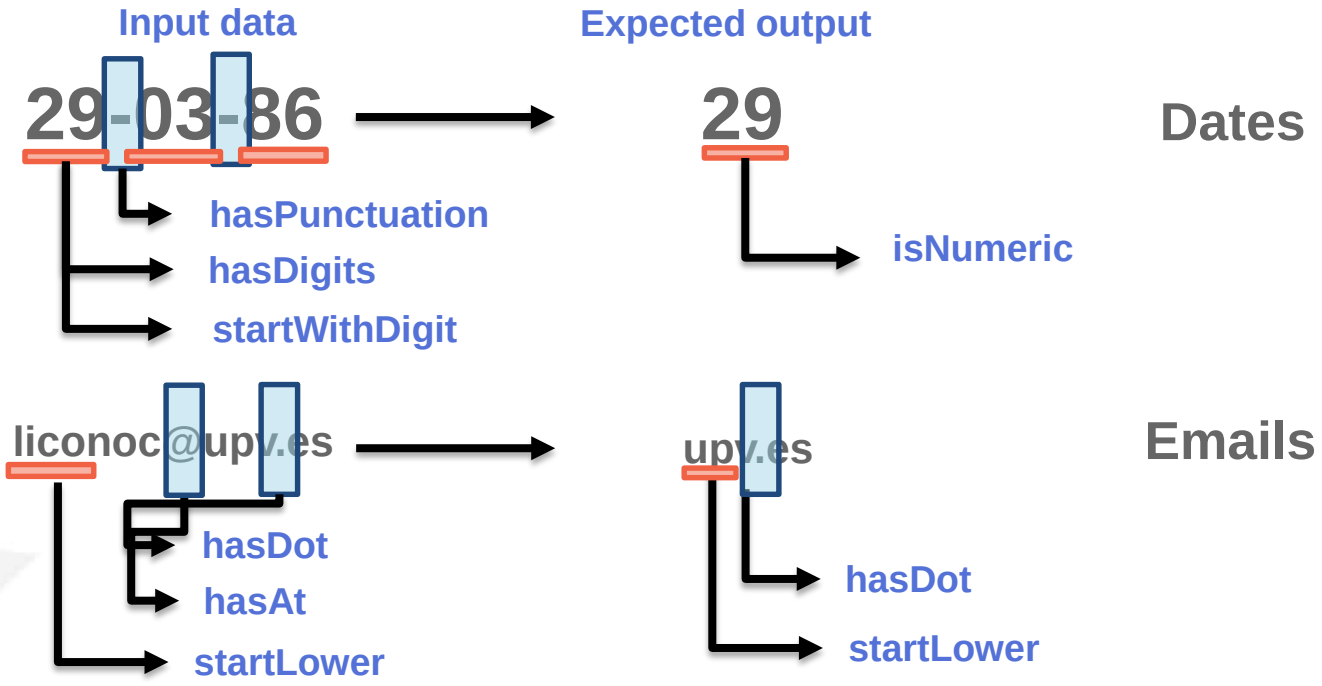
2. Dynamic Background Knowledge

Detecting the domain/problem

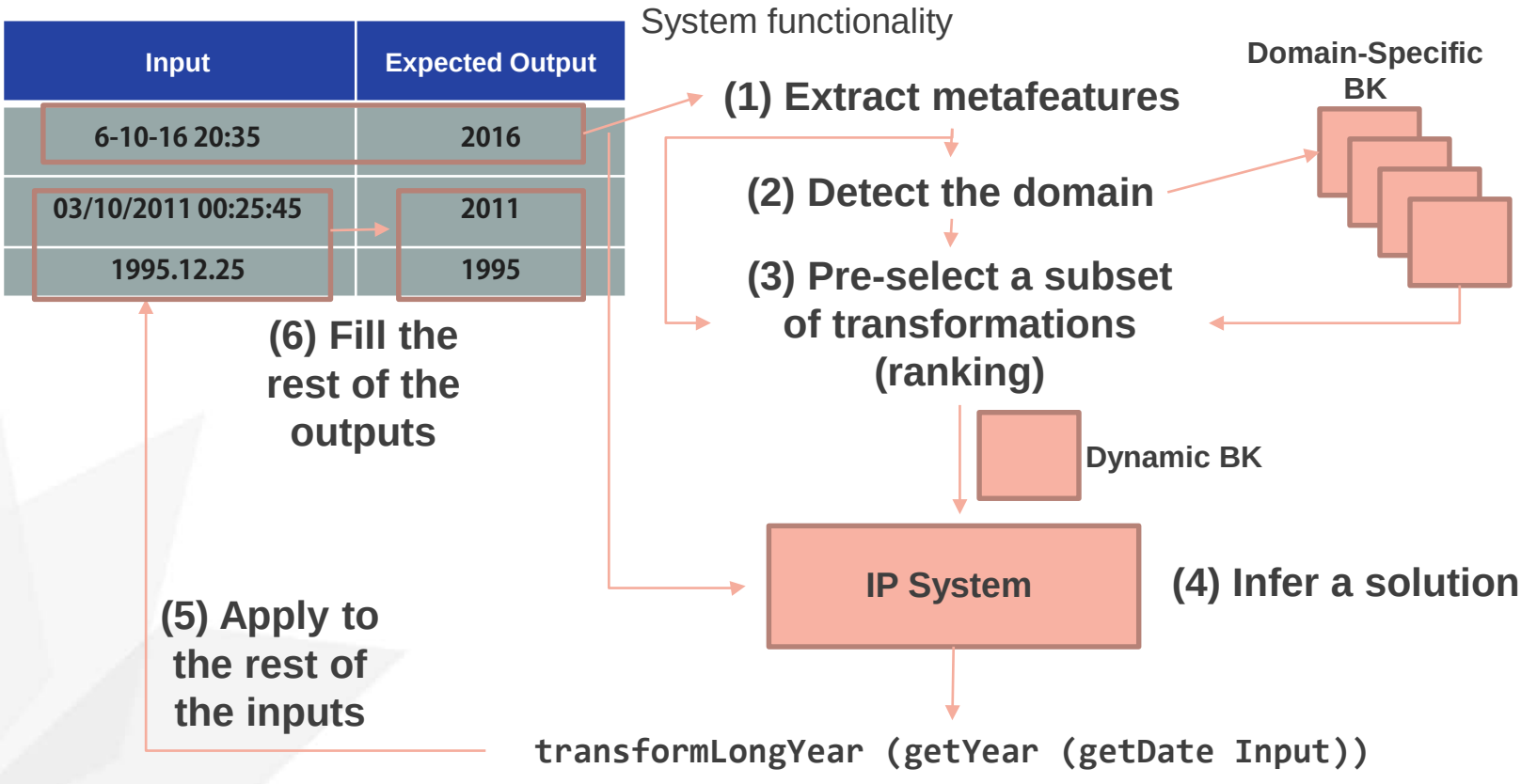
Metafeatures

Useful for:

- Detecting the domain
- Detecting the problem



2. Dynamic Background Knowledge



Adaptive Domain Detection

Joint work with Gust Verbruggen, Luc De Raedt José Hernández Orallo and Cèsar Ferri

Work in progress

KU LEUVEN

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

dmip

DTAI
Declarative Languages & Artificial Intelligence

3. Adaptive Domain Detection

Work in progress

Adrián Palacios Corella	apalacios@dsic.upv.es	C/ José Todolí Cucarella, 22	03/04/17 19:39
Cloud Bouker	cle@hormail.com	Rua borg, 115	27/06/2017 22h56
Mr David Meza	963347458	4033 Lovers Lane Dickinson, Texas 77539	10/2/2017 12:30 PM

Recognise and update known domains

Detect errors

Addresses

Detect and learn new domains

Dates

Times

Dates	Times
03/04/17	19:39
27/06/2017	22h56
10/2/2017	12:30 PM

Split examples using the domains

Thank you!

liconoc@upv.es
@liconoc