Temporal Alignment

University of Zurich[UZH]

Anton Dignös, Michael H. Böhlen, and Johann Gamper

{adignoes, boehlen}@ifi.uzh.ch, gamper@inf.unibz.it

Freie Universität Bozen
Libera Università di Bolzano
Free University of Bozen · Bolzano

Goal and Approach

Goal: Reduction of temporal operators to nontemporal operators using adjustment of timestamps.

Problem Definition: Given a temporal operator ψ^T, and input relations $\mathbf{r}_1, \dots \mathbf{r}_n$, our goal is to express $\psi^T(\mathbf{r}_1, \dots \mathbf{r}_n)$ as follows:

$$\psi^T\Big(\mathbf{r}_1, \dots \mathbf{r}_n\Big) = \psi\Big(\mathcal{P}^T(\mathbf{r}_1, \dots \mathbf{r}_n), \dots \mathcal{P}^T(\mathbf{r}_n, \dots \mathbf{r}_1)\Big) \qquad \textbf{(reduction)}$$

where ψ is the nontemporal operator corresponding to ψ^T, and $\mathcal{P}^T$ is a temporal primitive.

Solution:

- Two new algebra operators (primitives) for the adjustment of timestamps:
  - Temporal Splitter $\mathcal{N}$
  - Temporal Aligner ϕ
- Reduction rules for usage within nontemporal RA.
- Timestamp propagation for accessing original timestamps.

Key Points

Reduction rules that satisfy three key properties:

- Reducible to nontemporal queries on each snapshot.
  - $\forall t : \tau_t(\psi^T(\mathbf{D^T})) \equiv \psi(\tau_t(\mathbf{D^T}))$
- Original Timestamps are accessible.
  - $\forall t : \tau_t(\psi^T(\mathbf{D^T})) \equiv \psi(\tau_t(\epsilon(\mathbf{D^T})))$
- Interval boundaries of input are preserved.
  1. $\forall t \in z.T : \mathsf{L}[\psi^T(\mathbf{D^T})](z,t)$ is equal
  2. $z.T$ is maximal with respect to 1

where τ_t is the timeslice operator, ϵ propagates original timestamps, and $\mathsf{L}[\psi^T(\mathbf{D^T})](z,t)$ is the lineage set of result tuple z for $\psi^T(\mathbf{D^T})$ at time point t.

Example

Input: Employee N employed at department D during time T.

R

	N	D	T
r_1	Joe	DB	[Feb, Jul)
r_2	Ann	DB	[Feb, Sep)
r_3	Sam	AI	[May, Oct)

Query: What is the time-varying average duration of contracts per department?

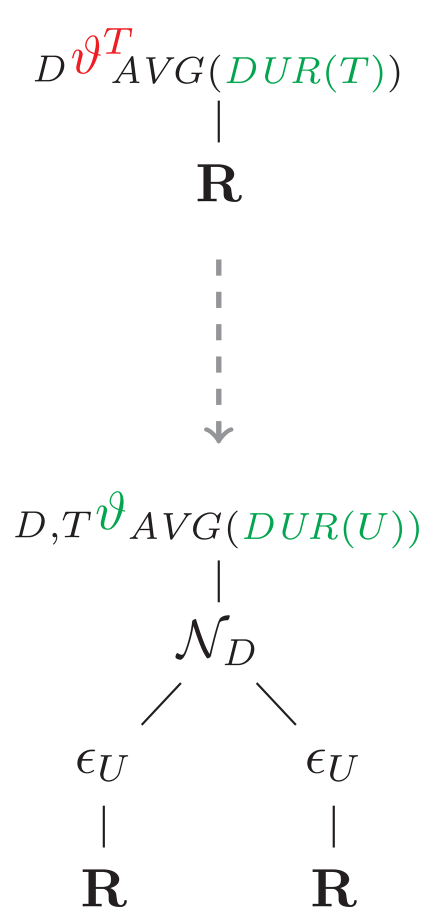
Result: Temporal Aggregation $_D\vartheta^T_{AVG(DUR(T))}(\mathbf{R})$

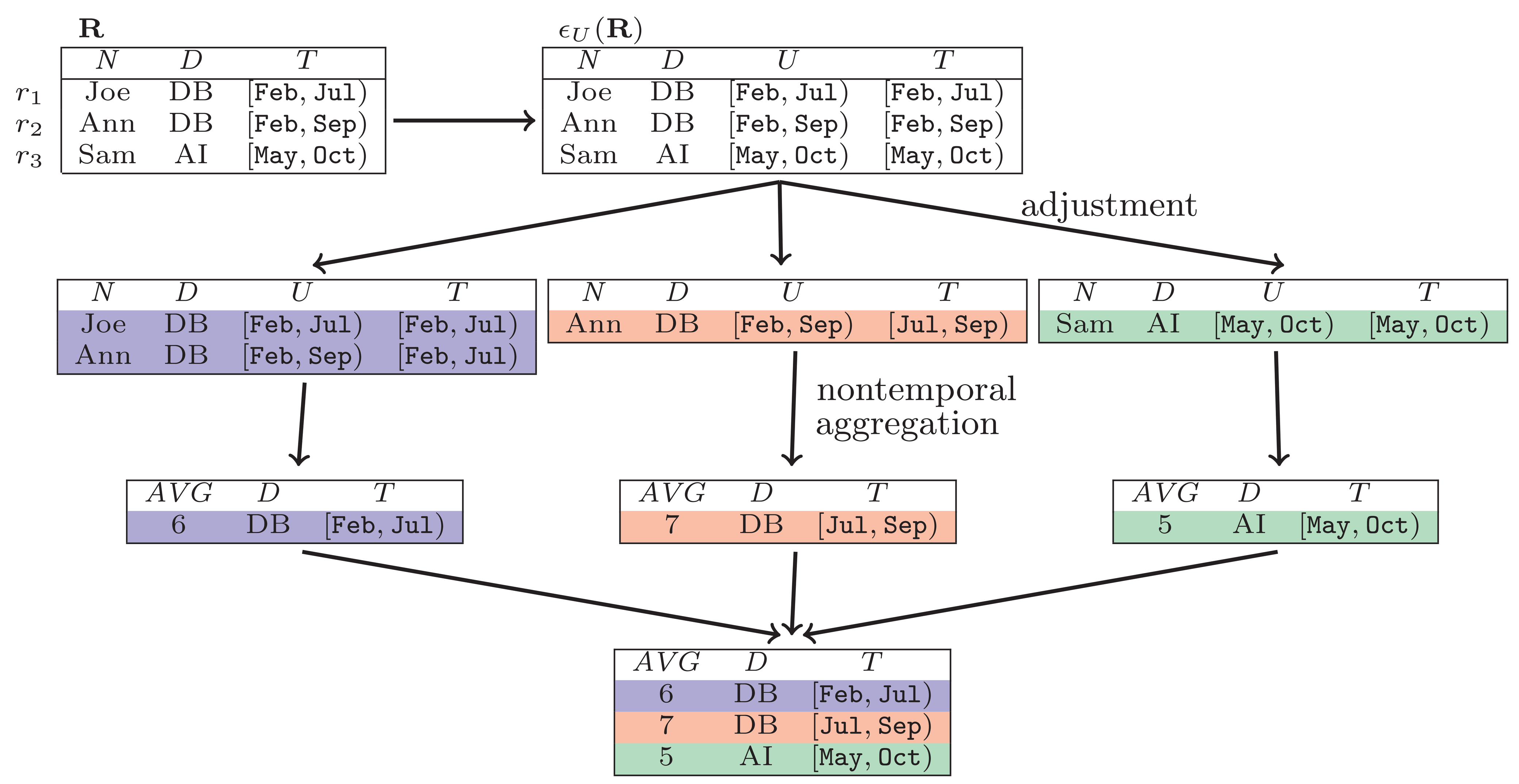
	AVG	D	T
z_1	6	DB	[Feb, Jul)
z_2	7	DB	[Jul, Sep)
z_3	5	AI	[May, Oct)

Processing steps:

Query: $_D\vartheta^T_{AVG(DUR(T))}(\mathbf{R})$

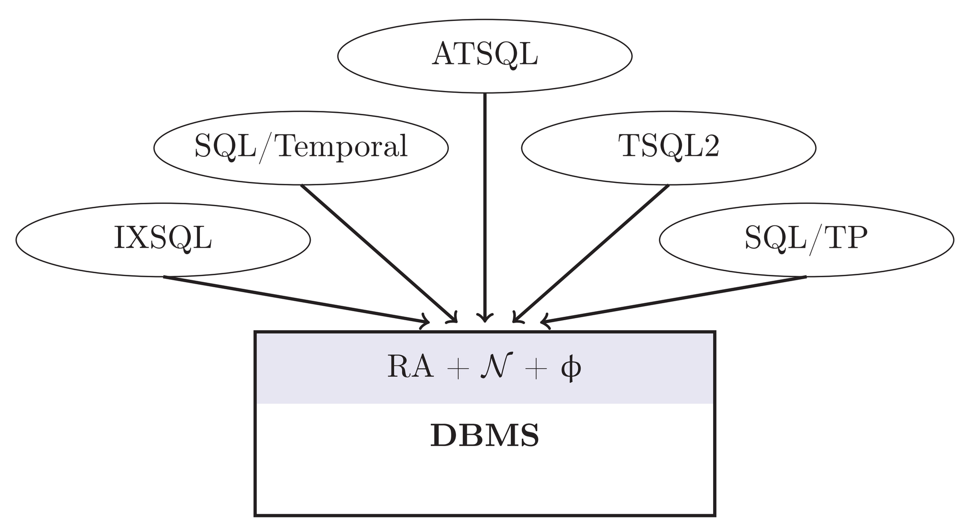
1. Timestamp propagation: $_D\vartheta^T_{AVG(DUR(T))}(\epsilon_U(\mathbf{R}))$
2. Timestamp substitution: $_D\vartheta^T_{AVG(DUR(U))}(\epsilon_U(\mathbf{R}))$
3. Temporal adjustment: $\mathbf{R}' \leftarrow \mathcal{N}_D(\epsilon_U(\mathbf{R}), \epsilon_U(\mathbf{R}))$
4. Nontemporal aggregation: $_{D,T}\vartheta_{AVG(DUR(U))}(\mathbf{R}')$





Implementation

Operator		Reduction
$\sigma^T_\theta(\mathbf{r})$	$=$	$\sigma_\theta(\mathbf{r})$
$\pi^T_\mathbf{B}(\mathbf{r})$	$=$	$\pi_{\mathbf{B},T}(\mathcal{N}_\mathbf{B}(\mathbf{r},\mathbf{r}))$
$_\mathbf{B}\vartheta^T_F(\mathbf{r})$	$=$	$_{\mathbf{B},T}\vartheta_F(\mathcal{N}_\mathbf{B}(\mathbf{r},\mathbf{r}))$
$\mathbf{r} -^T \mathbf{s}$	$=$	$\mathcal{N}_\mathbf{A}(\mathbf{r},\mathbf{s}) - \mathcal{N}_\mathbf{A}(\mathbf{s},\mathbf{r})$
$\mathbf{r} \cup^T \mathbf{s}$	$=$	$\mathcal{N}_\mathbf{A}(\mathbf{r},\mathbf{s}) \cup \mathcal{N}_\mathbf{A}(\mathbf{s},\mathbf{r})$
$\mathbf{r} \cap^T \mathbf{s}$	$=$	$\mathcal{N}_\mathbf{A}(\mathbf{r},\mathbf{s}) \cap \mathcal{N}_\mathbf{A}(\mathbf{s},\mathbf{r})$
$\mathbf{r} \times^T \mathbf{s}$	$=$	$\alpha(\phi_\top(\mathbf{r},\mathbf{s}) \bowtie_{\mathbf{r}.T=\mathbf{s}.T} \phi_\top(\mathbf{s},\mathbf{r}))$
$\mathbf{r} \bowtie^T_\theta \mathbf{s}$	$=$	$\alpha(\phi_\theta(\mathbf{r},\mathbf{s}) \bowtie_{\theta\wedge\mathbf{r}.T=\mathbf{s}.T} \phi_\theta(\mathbf{s},\mathbf{r}))$
$\mathbf{r}$ ⟖$^T_\theta$ $\mathbf{s}$	$=$	$\alpha(\phi_\theta(\mathbf{r},\mathbf{s})$ ⟖$_{\theta\wedge\mathbf{r}.T=\mathbf{s}.T}$ $\phi_\theta(\mathbf{s},\mathbf{r}))$
$\mathbf{r}$ ⟕$^T_\theta$ $\mathbf{s}$	$=$	$\alpha(\phi_\theta(\mathbf{r},\mathbf{s})$ ⟕$_{\theta\wedge\mathbf{r}.T=\mathbf{s}.T}$ $\phi_\theta(\mathbf{s},\mathbf{r}))$
$\mathbf{r}$ ⟗$^T_\theta$ $\mathbf{s}$	$=$	$\alpha(\phi_\theta(\mathbf{r},\mathbf{s})$ ⟗$_{\theta\wedge\mathbf{r}.T=\mathbf{s}.T}$ $\phi_\theta(\mathbf{s},\mathbf{r}))$
$\mathbf{r} \triangleright^T_\theta \mathbf{s}$	$=$	$\phi_\theta(\mathbf{r},\mathbf{s}) \triangleright_{\theta\wedge\mathbf{r}.T=\mathbf{s}.T} \phi_\theta(\mathbf{s},\mathbf{r})$



$\epsilon_U(\mathbf{r})$: `SELECT Ts Us, Te Ue, * FROM r`

$\mathcal{N}_\mathbf{B}(\mathbf{r},\mathbf{s})$: `FROM (r NORMALIZE s USING(B)) r`

$\phi_\theta(\mathbf{r},\mathbf{s})$: `FROM (r ALIGN s ON` θ`) r`

$\alpha(\mathbf{r})$: `SELECT ABSORB * FROM r`

http://www.ifi.uzh.ch/dbtg/research/align.html

Summary

- Algebraic basis for temporal operators.
- Reduction of temporal operators to nontemporal operators.
- Deep integration into PostgreSQL kernel.

Future Work

- Optimization/equivalence rules for temporal primitives.
- Extensions towards time depended (malleable) quantities.
- Extension to bag algebra.

Published in: *Proceedings of the 38th ACM SIGMOD International Conference on Management of Data*, Scottsdale, Arizona, USA, May 2012.