Incremental Discovery of Prominent Situational Facts

Afroza Sultana[1], Naeemul Hassan[1], Chengkai Li[1], Jun Yang[2], Cong Yu[3]
[1]University of Texas at Arlington, [2]Duke University, [3]Google Research

ICDE 2014, Chicago, IL

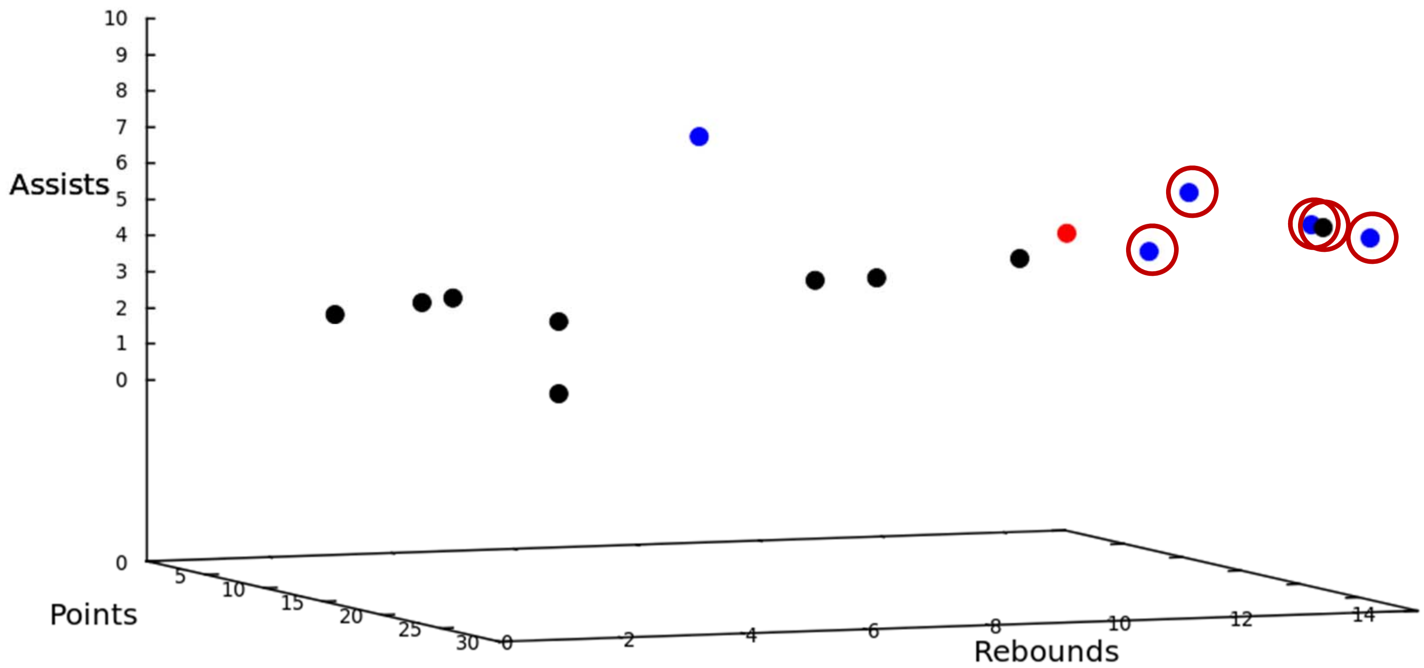
UNIVERSITY OF TEXAS ARLINGTON DUKE COMPUTER SCIENCE Google

Situational Facts

"Paul George had 21 points, 11 rebounds and 5 assists to become the first Pacers player with a 20/10/5 (points/rebounds/assists) game against the Bulls since Detlef Schrempf in December 1992."
(http://espn.go.com/espn/elias?date=20130205)

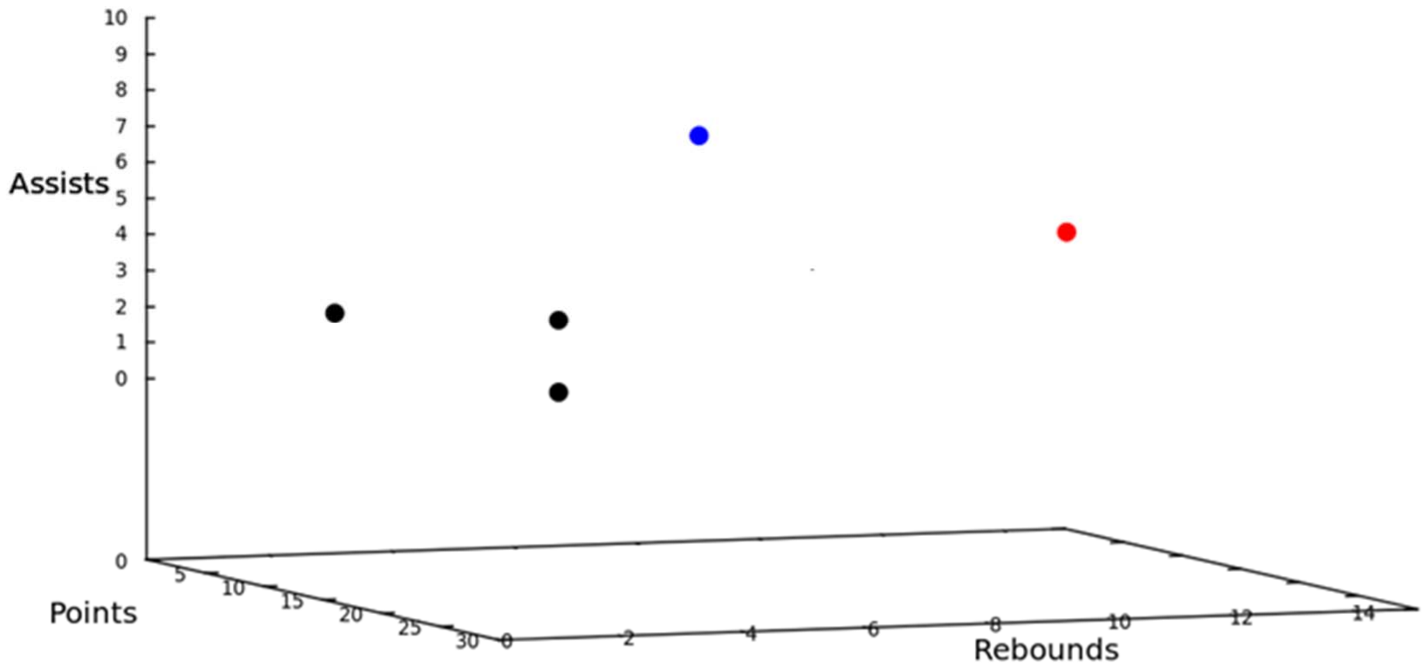
Situational Facts

"Paul George had 21 points, 11 rebounds and 5 assists to become the first Pacers player with a 20/10/5 (points/rebounds/assists) game against the Bulls since Detlef Schrempf in December 1992."
(http://espn.go.com/espn/elias?date=20130205)



Situational Facts

"Paul George had 21 points, 11 rebounds and 5 assists to become the first Pacers player with a 20/10/5 (points/rebounds/assists) game against the Bulls since Detlef Schrempf in December 1992."
(http://espn.go.com/espn/elias?date=20130205)



Situational Facts

“The social world’s most viral photo ever generated 3.5 million likes, 170,000 comments and 460,000 shares by Wednesday afternoon.”

(http://www.cnbc.com/id/49728455/President Obama Sets New Social Media Record)

Situational Facts

“The social world’s most viral photo ever generated 3.5 million likes, 170,000 comments and 460,000 shares by Wednesday afternoon.”

(http://www.cnbc.com/id/49728455/President Obama Sets New Social Media Record)

Situational Facts

"The social world's most viral photo ever generated 3.5 million likes, 170,000 comments and 460,000 shares by Wednesday afternoon."

(http://www.cnbc.com/id/49728455/President Obama Sets New Social Media Record)

Situational Facts

- Stock Data: Stock A becomes the first stock in history with price over $300 and market cap over $400 billion.

- Weather Data: Today's measures of wind speed and humidity are x and y, respectively. City B has never encountered such high wind speed and humidity in March.

- Criminal Records: There were 50 DUI arrests and 20 collisions in city C yesterday, the first time in 2013.

Financial Analyst
Journalists
Scientists
Citizens

A Mini-world of Basketball Gamelogs

id	player	day	month	season	team	opp_team	pts	ast	reb
t_1	Bogues	11	Feb.	1991-92	Hornets	Hawks	4	12	5
t_2	Seikaly	13	Feb.	1991-92	Heat	Hawks	24	5	15
t_3	Sherman	7	Dec.	1993-94	Celtics	Nets	13	13	5
t_4	Wesley	4	Feb.	1994-95	Celtics	Nets	2	5	2
t_5	Wesley	5	Feb.	1994-95	Celtics	Timberwolves	3	5	3
t_6	Strictland	3	Jan.	1995-96	Blazers	Celtics	27	18	8
t_7	Wesley	25	Feb.	1995-96	Celtics	Nets	12	13	5

Last tuple appended to table

A Mini-world of Basketball Gamelogs

id	player	day	month	season	team	opp_team	pts	ast	reb
t_1	Bogues	11	Feb.	1991-92	Hornets	Hawks	4	12	5
t_2	Seikaly	13	Feb.	1991-92	Heat	Hawks	24	5	15
t_3	Sherman	7	Dec.	1993-94	Celtics	Nets	13	13	5
t_4	Wesley	4	Feb.	1994-95	Celtics	Nets	2	5	2
t_5	Wesley	5	Feb.	1994-95	Celtics	Timberwolves	3	5	3
t_6	Strictland	3	Jan.	1995-96	Blazers	Celtics	27	18	8
t_7	Wesley	25	Feb.	1995-96	Celtics	Nets	12	13	5

A Mini-world of Basketball Gamelogs

id			month				pts	ast	reb
t_1			Feb.				4	12	5
t_2			Feb.				24	5	15
t_4			Feb.				2	5	2
t_5			Feb.				3	5	3
t_7			Feb.				12	13	5

A Mini-world of Basketball Gamelogs

id			month				pts	ast	reb
t_1			Feb.				4	12	5
t_2			Feb.				24	5	15
t_4			Feb.				2	5	2
t_5			Feb.				3	5	3
t_7			Feb.				12	13	5

- Wesley had 12 points, 13 assists and 5 rebounds on February 25, 1996 to become the first player with a 12/13/5 (points/assists/rebounds) in February.

A Mini-world of Basketball Gamelogs

id				**season**			**pts**	**ast**	**reb**
t_6				1995-96			27	18	8
t_7				1995-96			12	13	5

A Mini-world of Basketball Gamelogs

id					team	opp_team		ast	reb
t_3					Celtics	Nets		13	5
t_4					Celtics	Nets		5	2
t_7					Celtics	Nets		13	5

- Wesley had 13 assists and 5 rebounds on February 25, 1996 to become the second Celtics player with a 13/5 (assists/rebounds) game against the Nets.

Problem Definition

Dimension space: $\mathcal{D}=\{d_1,\ldots,d_n\}$

Measure space: $\mathcal{M}=\{m_1,\ldots,m_s\}$

id	player	day	month	season	team	opp_team	pts	ast	reb
t_1	Bogues	11	Feb.	1991-92	Hornets	Hawks	4	12	5
t_2	Seikaly	13	Feb.	1991-92	Heat	Hawks	24	5	15
t_3	Sherman	7	Dec.	1993-94	Celtics	Nets	13	13	5
t_4	Wesley	4	Feb.	1994-95	Celtics	Nets	2	5	2
t_5	Wesley	5	Feb.	1994-95	Celtics	Timberwolves	3	5	3
t_6	Strictland	3	Jan.	1995-96	Blazers	Celtics	27	18	8

append-only table

Problem Definition

- Constraint (C): $d_1=v_1 \wedge d_2=v_2 \wedge . . . \wedge\, d_n=v_n,\ v_i \in dom(d_i) \cup \{*\}$
  - team=*Celtics* ∧ opp_team=*Nets*

id					team	opp_team			
t_3					Celtics	Nets			
t_4					Celtics	Nets			

Problem Definition

- Constraint-Measure Pair (C, M): Combination of a constraint and measure subspace
  - (team=*Celtics* ∧ opp_team=*Nets*, {assists,rebounds})

id					team	opp_team		ast	reb
t_3					Celtics	Nets		13	5
t_4					Celtics	Nets		5	2

Problem Definition

- **Contextual skyline:** skyline regarding (C, M)
  - $\sigma_{\text{team}=\mathit{Celtics}\ \wedge\ \text{opp\_team}=\mathit{Nets}}(R)$, M={assists,rebounds}
    - $\{t_3\}$

id					team	opp_team		ast	reb
t_3					Celtics	Nets		13	5
t_4					Celtics	Nets		5	2

Problem Definition; Situational Fact Discover Problem

Tuples capturing real world events appended to table

id	player	day	month	season	team	opp_team	pts	ast	reb
t_1	Bogues	11	Feb.	1991-92	Hornets	Hawks	4	12	5
t_2	Seikaly	13	Feb.	1991-92	Heat	Hawks	24	5	15
t_3	Sherman	7	Dec.	1993-94	Celtics	Nets	13	13	5
t_4	Wesley	4	Feb.	1994-95	Celtics	Nets	2	5	2
t_5	Wesley	5	Feb.	1994-95	Celtics	Timberwolves	3	5	3
t_6	Strictland	3	Jan.	1995-96	Blazers	Celtics	27	18	8
t_7	Wesley	25	Feb.	1995-96	Celtics	Nets	12	13	5

Find constraint-measure pair (C,M) such that t is in the contextual skyline.

Constraint	Measure
month=*Feb*	pts, ast, rb
opp_team=*Nets*	ast, rb
team=*Celtics* ∧ opp_team=*Nets*	ast, rb
…	…

Template

Wesley had 12 points, 13 assists and 5 rebounds on February 25, 1996 to become the first player with a 12/13/5 (points/assists/rebounds) in February.

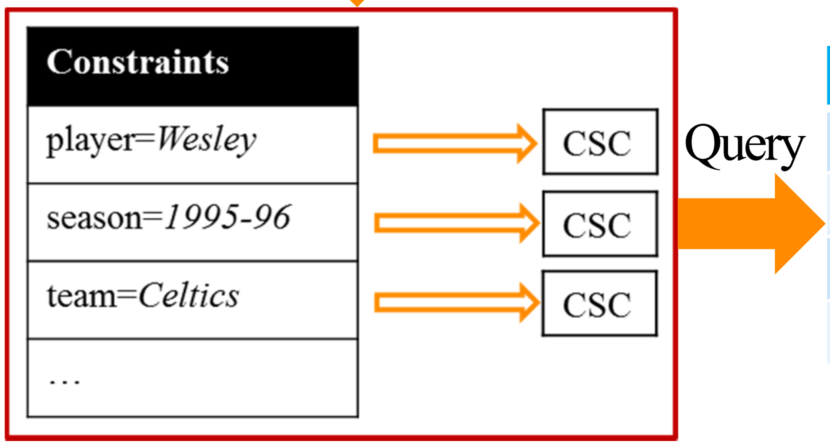
Related Work

- Conventional skyline analysis (Borzsonyi et al. ICDE 2001)
  - Q: context, measure subspace ⟹ A: contextual skyline tuples
    - ✓Our focus--- A: tuple ⟹ Q: constraint-measure pairs

Related Works

- Compressed Skycube (Xia et al. SIGMOD 2006)
  - Update compressed skycube in monitoring fashion
    - ✓We adapted CSC for each constraint: Constraint-CSC

id	player	day	month	season	team	opp_team	pts	ast	reb
t_1	Bogues	11	Feb.	1991-92	Hornets	Hawks	4	12	5
t_2	Seikaly	13	Feb.	1991-92	Heat	Hawks	24	5	15
t_3	Sherman	7	Dec.	1993-94	Celtics	Nets	13	13	5
t_4	Wesley	4	Feb.	1994-95	Celtics	Nets	2	5	2
t_5	Wesley	5	Feb.	1994-95	Celtics	Timberwolves	3	5	3
t_6	Strictland	3	Jan.	1995-96	Blazers	Celtics	27	18	8
t_7	Wesley	25	Feb.	1995-96	Celtics	Nets	12	13	5



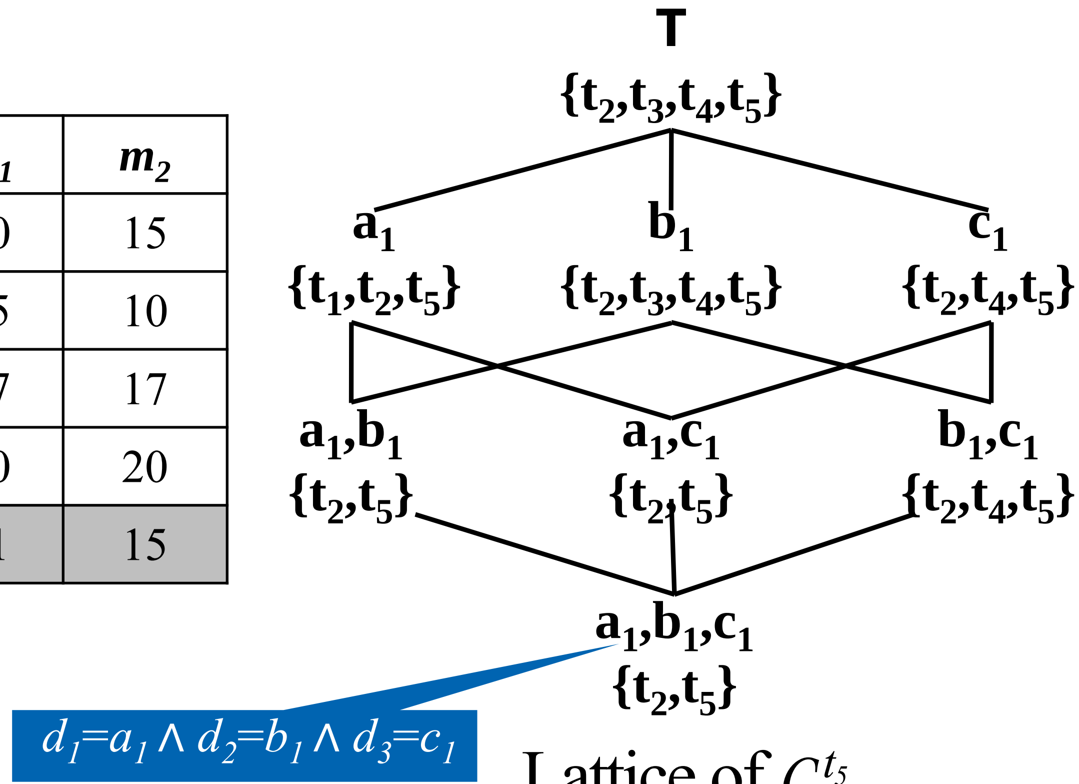
Constraint	Measure
month=*Feb*	pts, ast, rb
opp_team=*Nets*	ast, rb
team=*Celtics* ∧ opp_team=*Nets*	ast, rb
…	…

Related Works

- Prominent Analysis by Ranking (Wu et. Al. VLDB 2009)
  - Static data, onetime query
    - ✓We dealt on continuous data, standing query
  - Find the contexts where an object is ranked high in a single scoring attribute
    - ✓We considered skyline on multiple measure subspaces

Modeling

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15

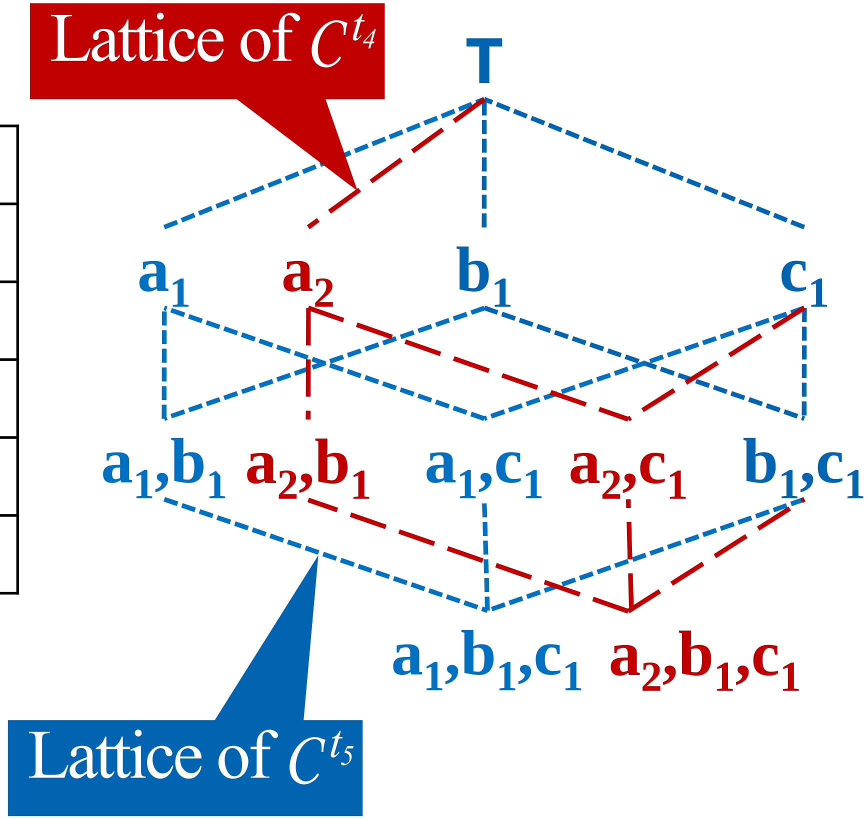


Lattice of C^{t_5}

Tuple Satisfied Constraint C^t: If $\forall d_i \in \mathcal{D}$, $C.d_i = *$ or $C.d_i = t.d_i$, t satisfies C.

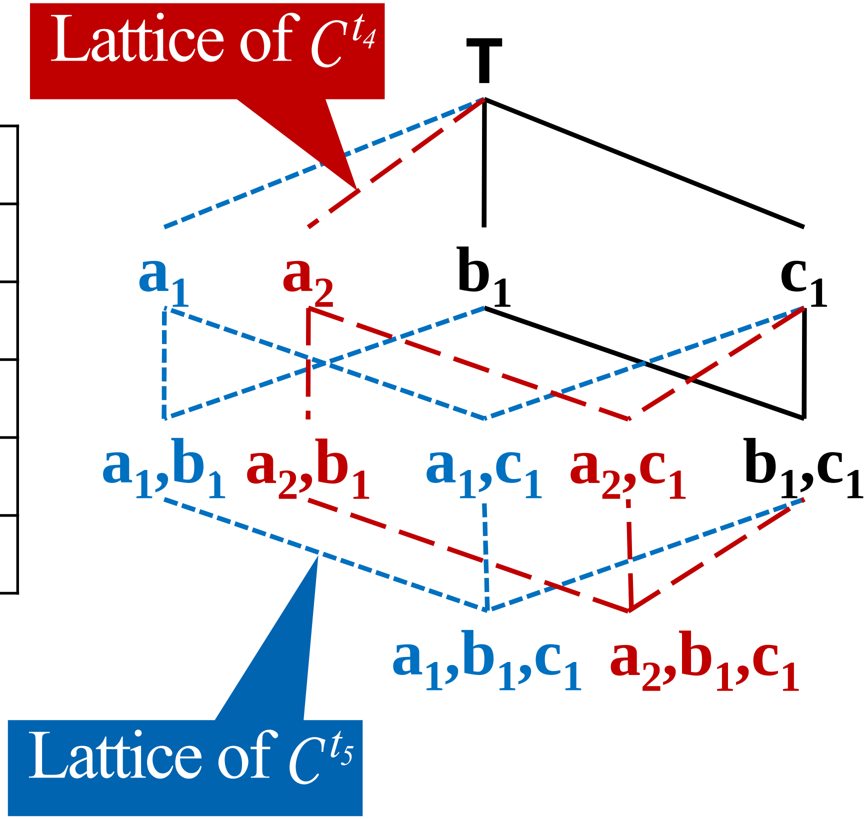
Modeling

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



Modeling

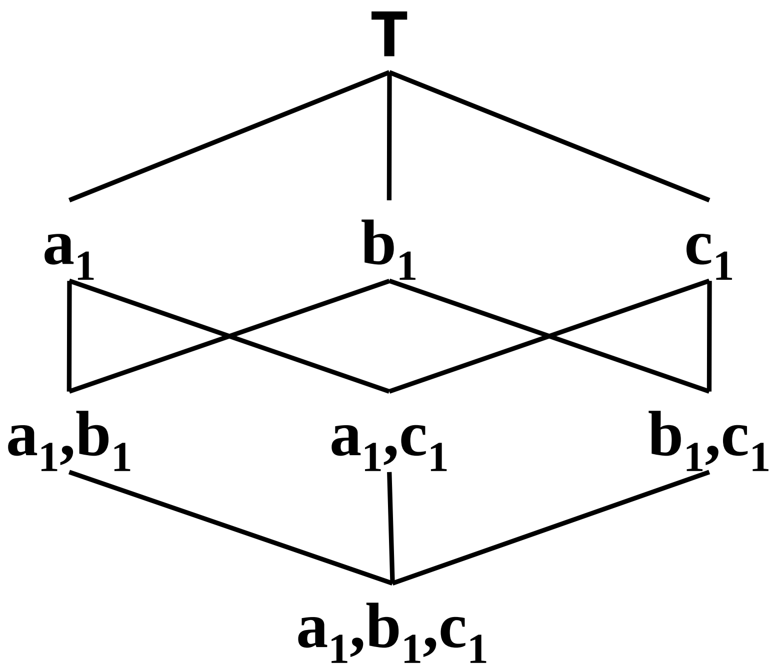
id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



Lattice Intersection: $C^{t_4,t_5}=C^{t_4} \cap C^{t_5}$

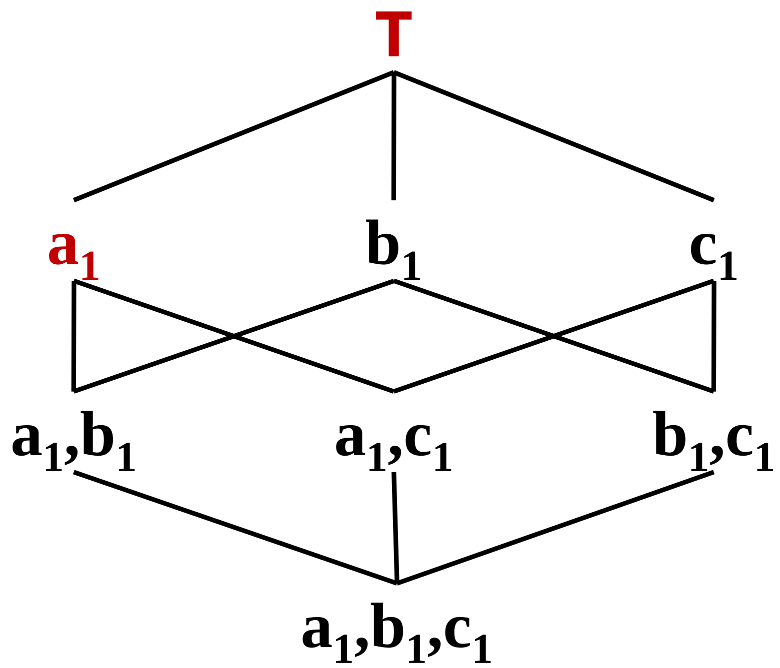
Brute-Force Approach

id	***d_1***	***d_2***	***d_3***	***m_1***	***m_2***
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



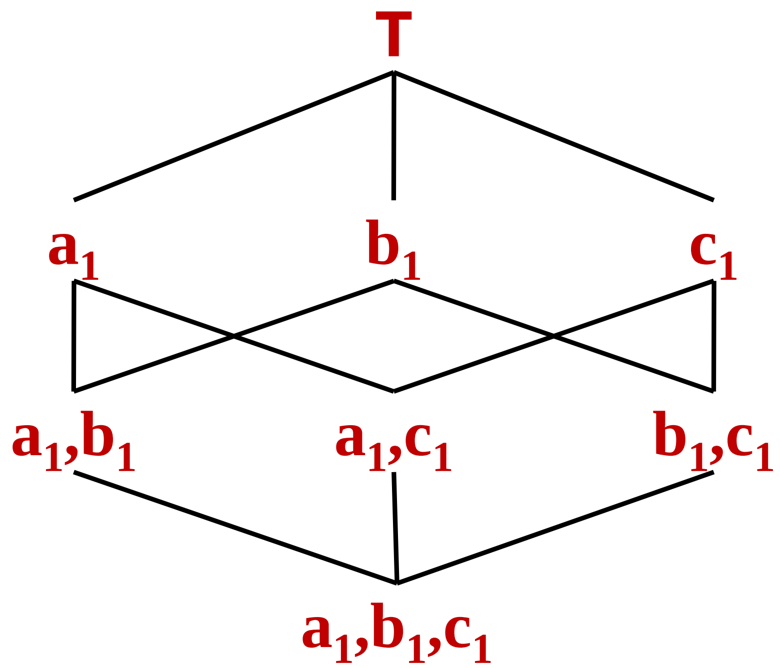
Brute-Force Approach

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



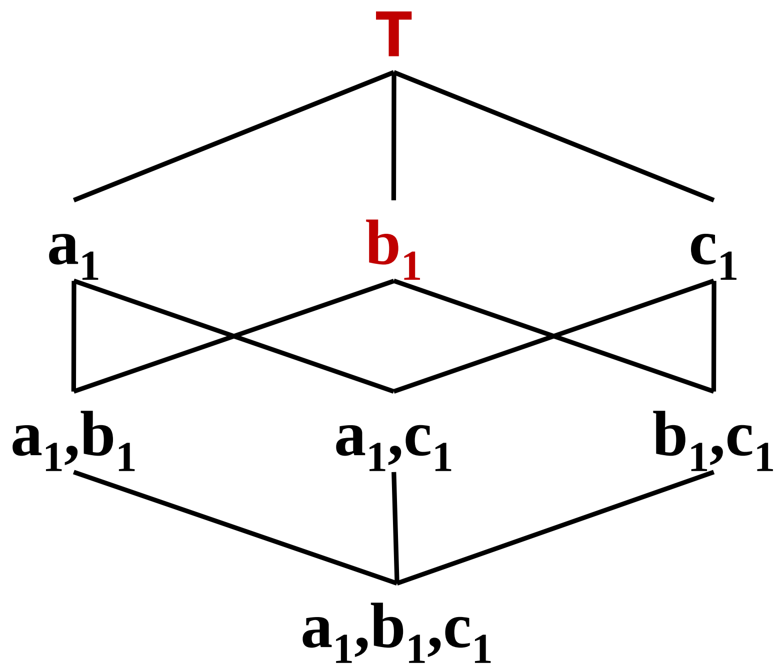
Brute-Force Approach

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



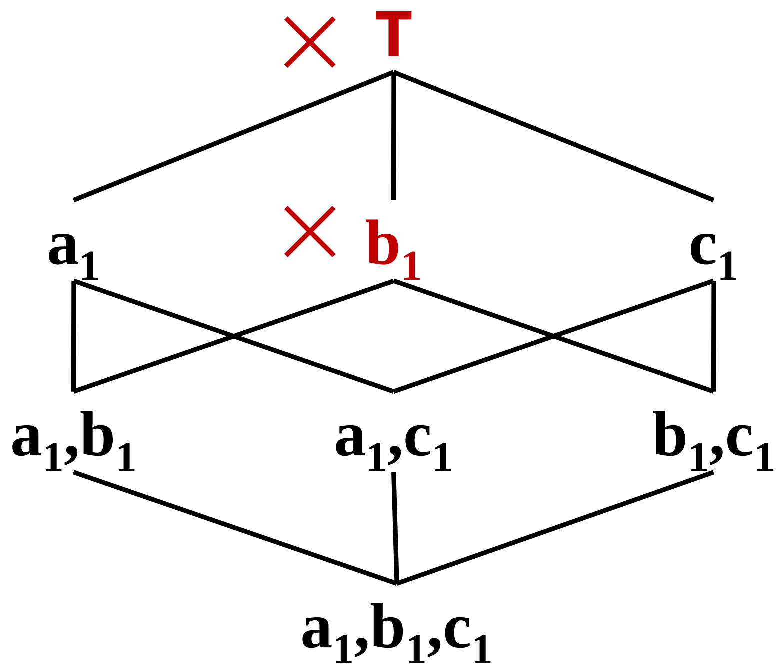
Brute-Force Approach

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



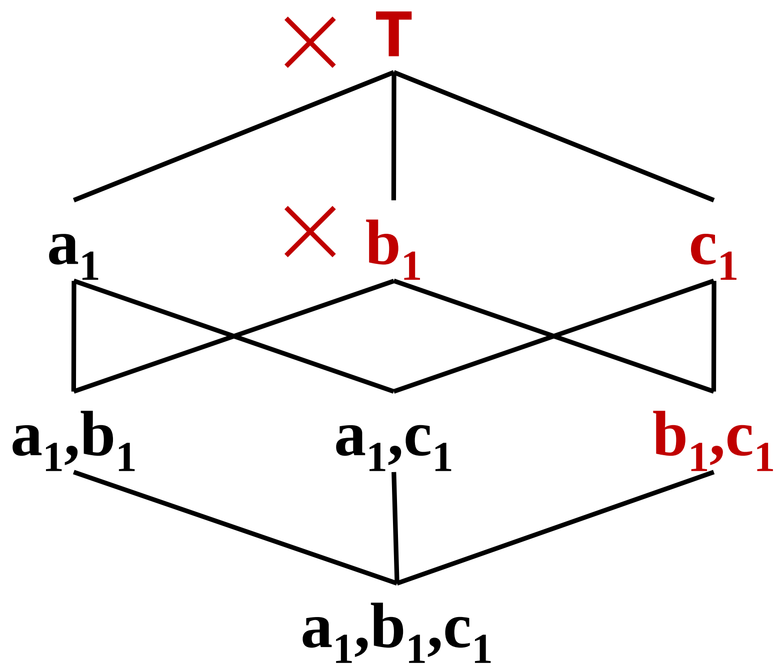
Brute-Force Approach

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



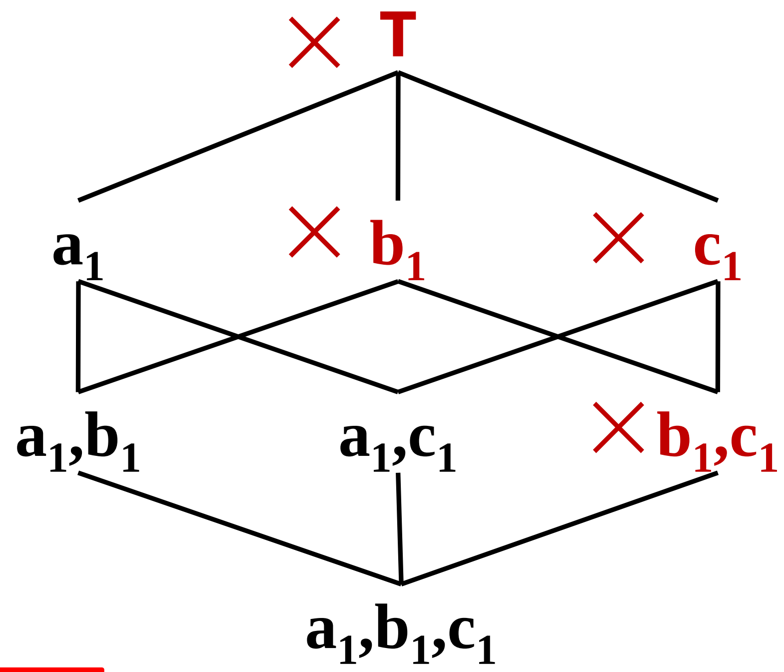
Brute-Force Approach

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



Brute-Force Approach

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



Total $|R|*(2^{|\mathcal{D}|+|\mathcal{M}|}-1)$ comparisons!
Total 16 comparisons in this case!

Challenges

- Exhaustive comparison with every tuple
- Under every constraint
- Over every measure subspace

Challenges and Ideas

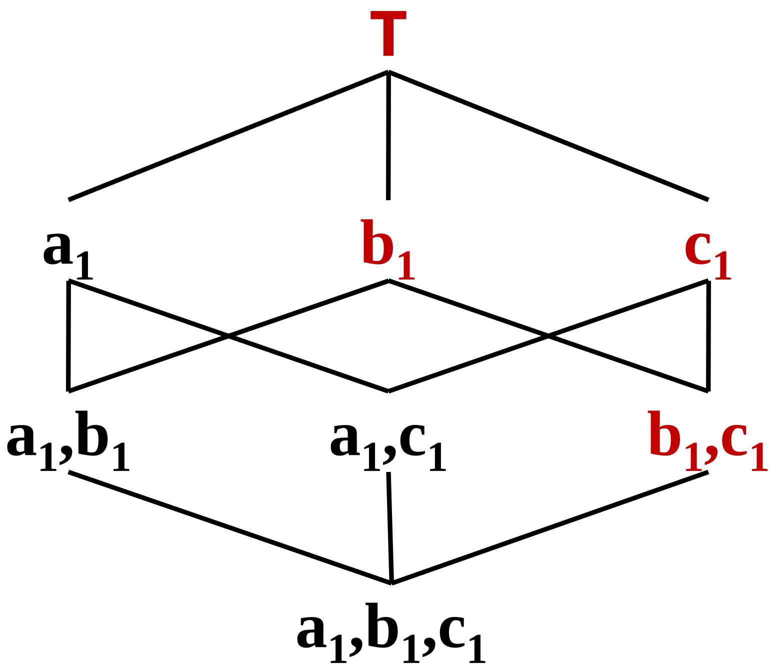
- Exhaustive comparison with every tuple
  - ✓Tuple reduction
    - Comparison with skyline tuples is enough
    - $t_4 \succ_{\{m_1,m_2\}} t_3 \succ_{\{m_1,m_2\}} t_5 \Rightarrow t_4 \succ_{\{m_1,m_2\}} t_5$

id		***d₂***		***m₁***	***m₂***
t_2		b_1		15	10
t_3		b_1		17	17
t_4		b_1		20	20
t_5		b_1		11	15

Challenges and Ideas

- Under every constraint
  - ✓ Constraint pruning
    - In $C^{t,t'}$, one comparison on t and t' is enough

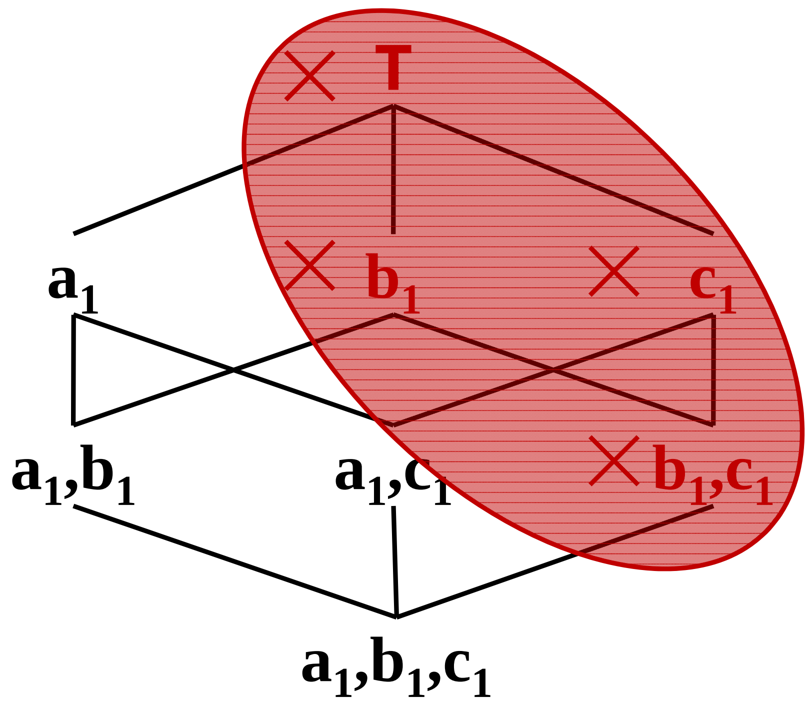
id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



Challenges and Ideas

- Under every constraint
  - ✓ Constraint pruning
    - In $C^{t,t'}$, one comparison on t and t' is enough

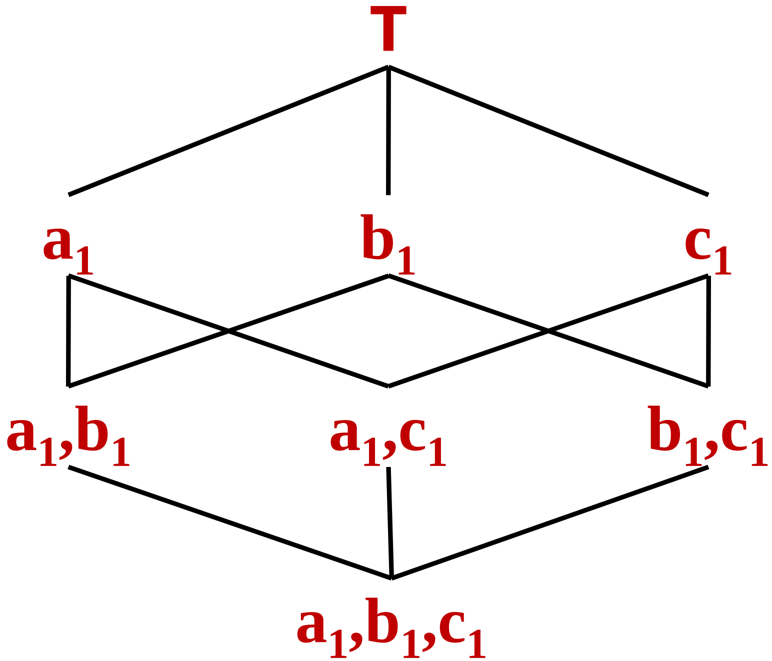
id	*d_1*	*d_2*	*d_3*	*m_1*	*m_2*
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



Challenges and Ideas

- Over every measure subspace
  - ✓Sharing computation across measure subspaces
    - Reusing computations on full space in subspaces

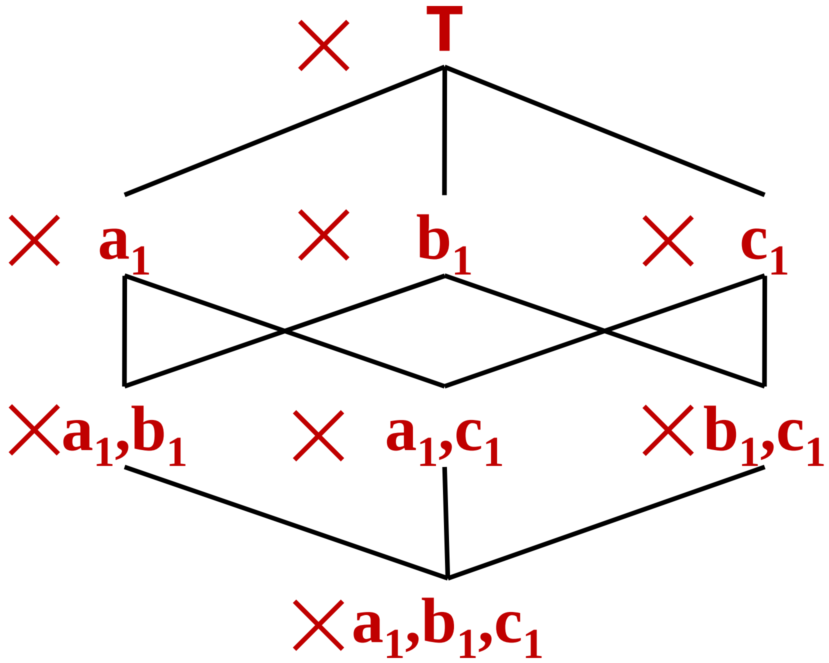
id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



Challenges and Ideas

- Over every measure subspace
  - ✓Sharing computation across measure subspaces
    - Reusing computations on full space in subspaces

id	d_1	d_2	d_3	m_1	
t_1	a_1	b_2	c_2	10	
t_2	a_1	b_1	c_1	15	
t_3	a_2	b_1	c_2	17	
t_4	a_2	b_1	c_1	20	
t_5	a_1	b_1	c_1	11	



Our Algorithms

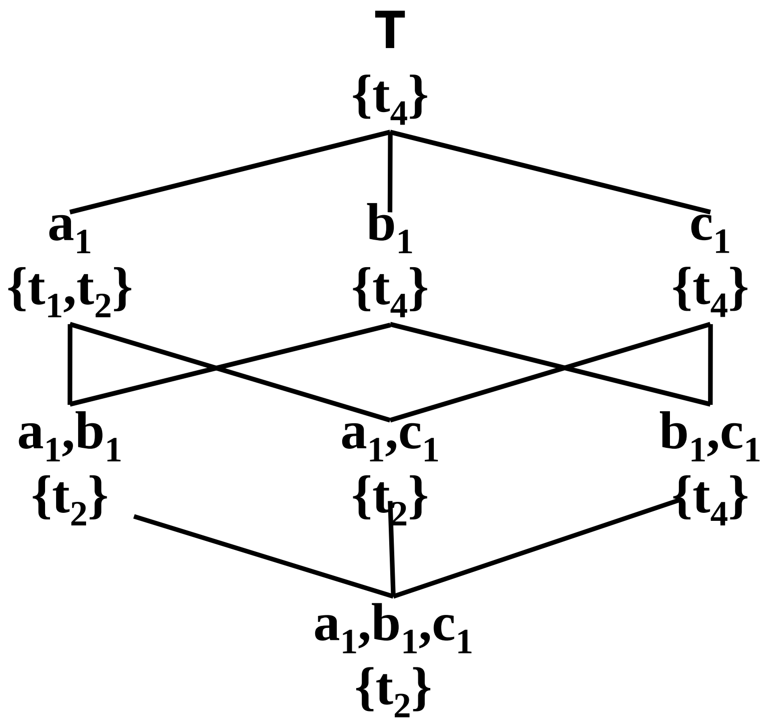
- Tuple reduction + Constraint pruning
    - BottomUp
    - TopDown
- Tuple reduction + Constraint pruning + Sharing computation
    - SBottomUp
    - STopDown

BottomUp

- Stores a tuple for every such constraint that qualifies it as a contextual skyline tuple
- Traverses the constraints in C^t in a bottom-up, breadth-first manner

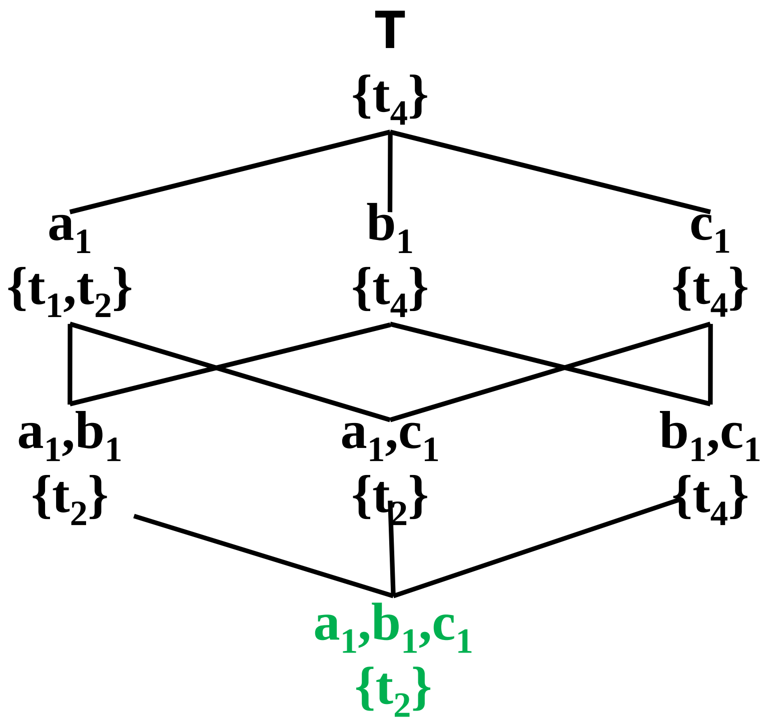
BottomUp

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



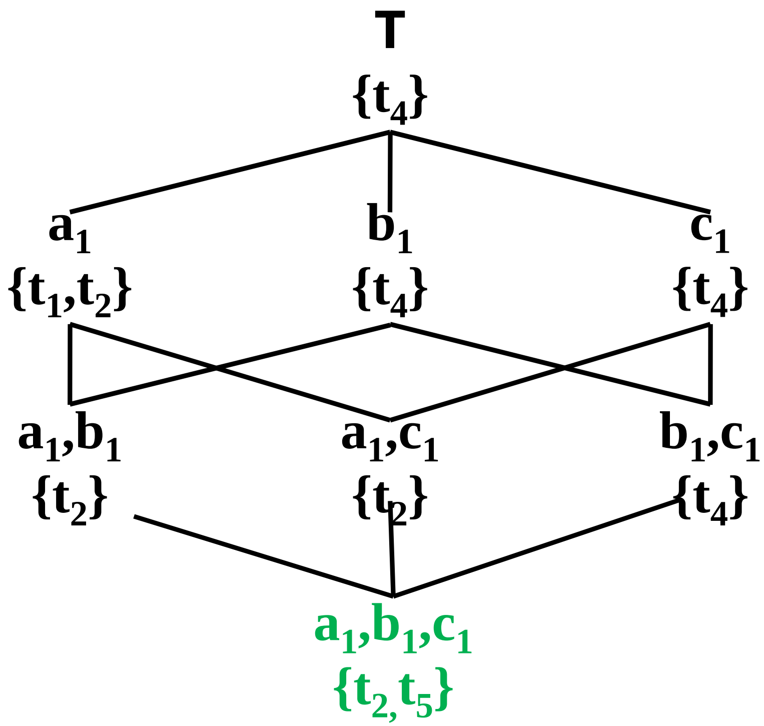
BottomUp

id	d_1	d_2	d_3	m_1	m_2
t_2	a_1	b_1	c_1	15	10
t_5	a_1	b_1	c_1	11	15



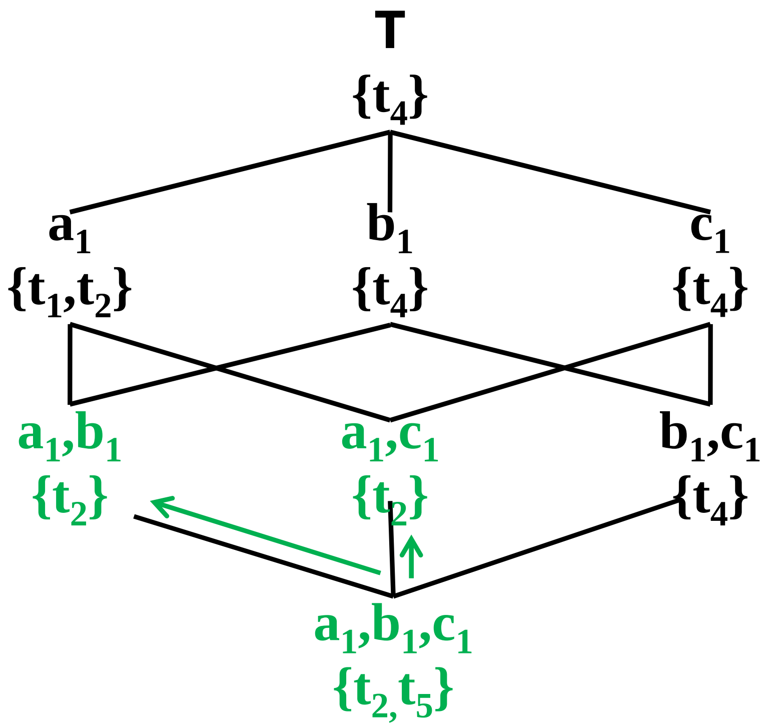
BottomUp

id	d_1	d_2	d_3	m_1	m_2
t_2	a_1	b_1	c_1	15	10
t_5	a_1	b_1	c_1	11	15



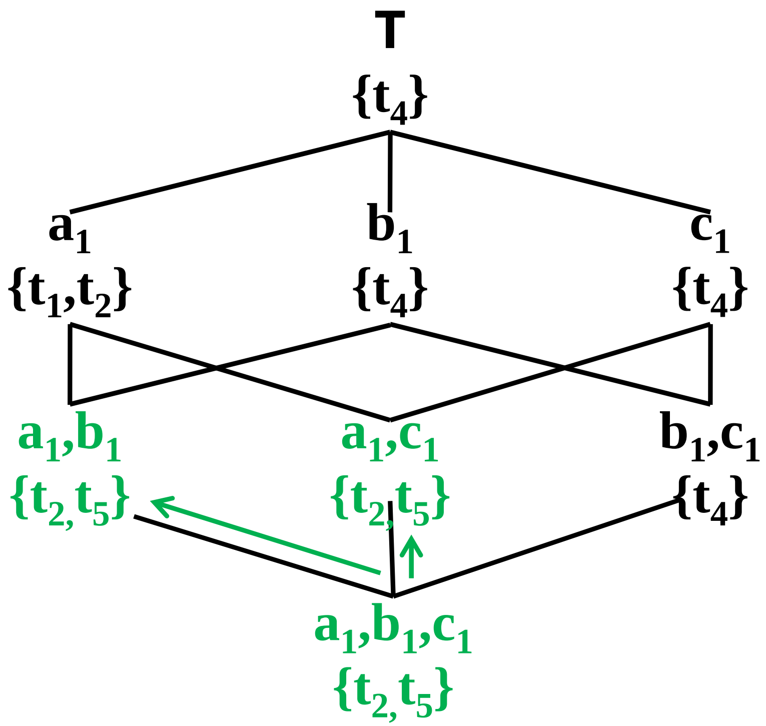
BottomUp

id	d_1	d_2	d_3	m_1	m_2
t_2	a_1	b_1	c_1	15	10
t_5	a_1	b_1	c_1	11	15



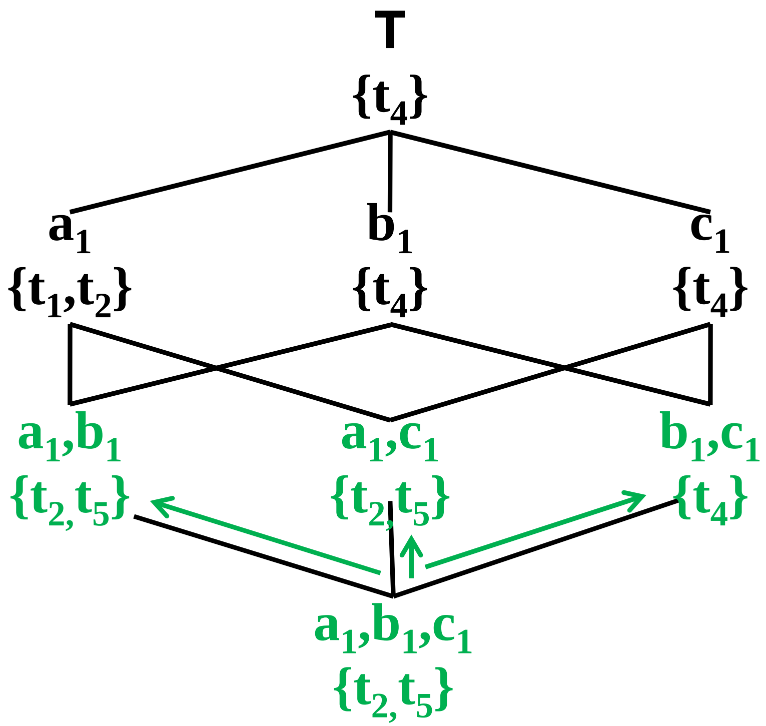
BottomUp

id	d_1	d_2	d_3	m_1	m_2
t_2	a_1	b_1	c_1	15	10
t_5	a_1	b_1	c_1	11	15



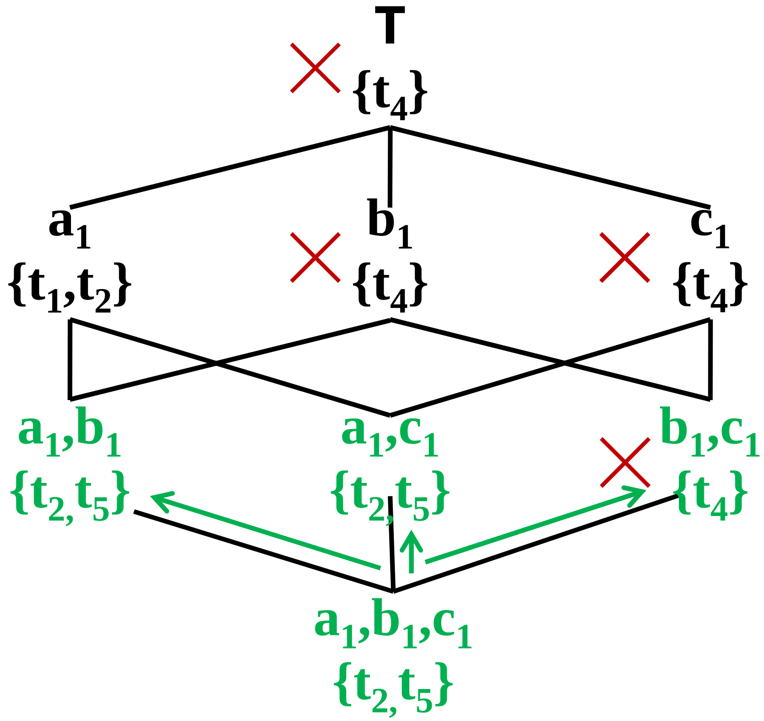
BottomUp

id		d_2	d_3	m_1	m_2
t_2		b_1	c_1	15	10
t_4		b_1	c_1	20	20
t_5		b_1	c_1	11	15



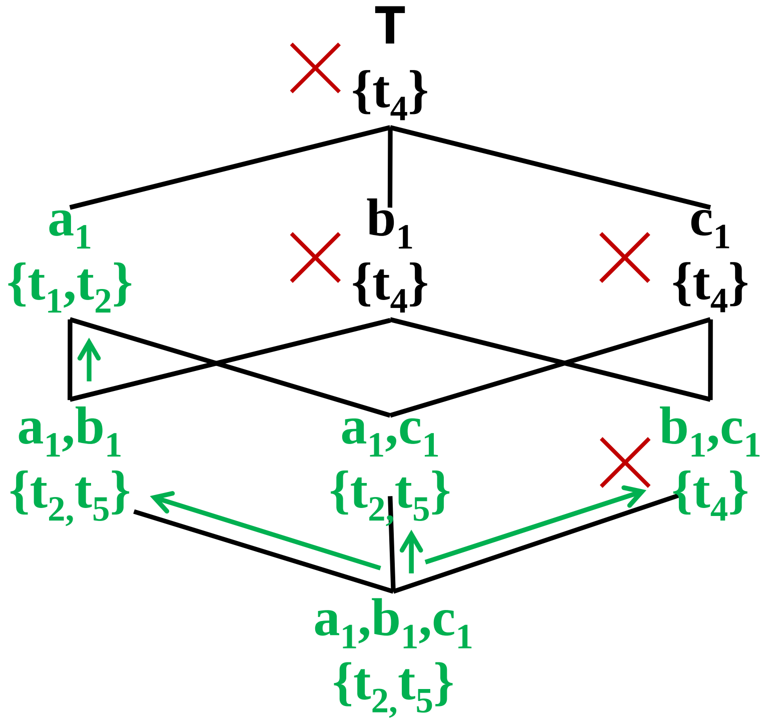
BottomUp

id		d_2	d_3	m_1	m_2
t_2		b_1	c_1	15	10
t_4		b_1	c_1	20	20
t_5		b_1	c_1	11	15



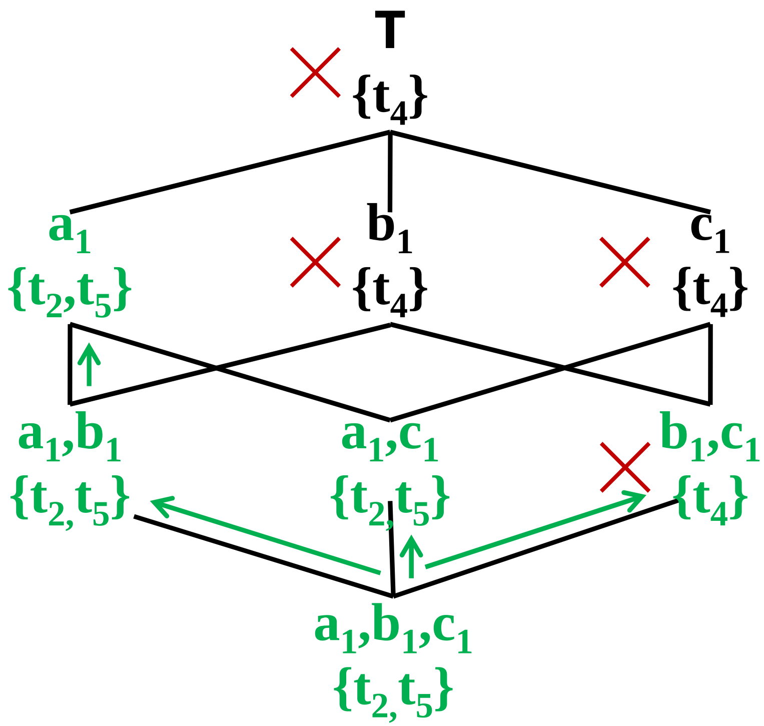
BottomUp

id	d_1			m_1	m_2
t_1	a_1			10	15
t_2	a_1			15	10
t_5	a_1			11	15



BottomUp

id	*d_1*			*m_1*	*m_2*
t_1	*a_1*			10	15
t_2	*a_1*			15	10
t_5	*a_1*			11	15



Total 6 comparisons in this case

BottomUp

- Cons of BottomUp
  - Repetitive storage: space complexity
  - Repetitive comparisons: time complexity

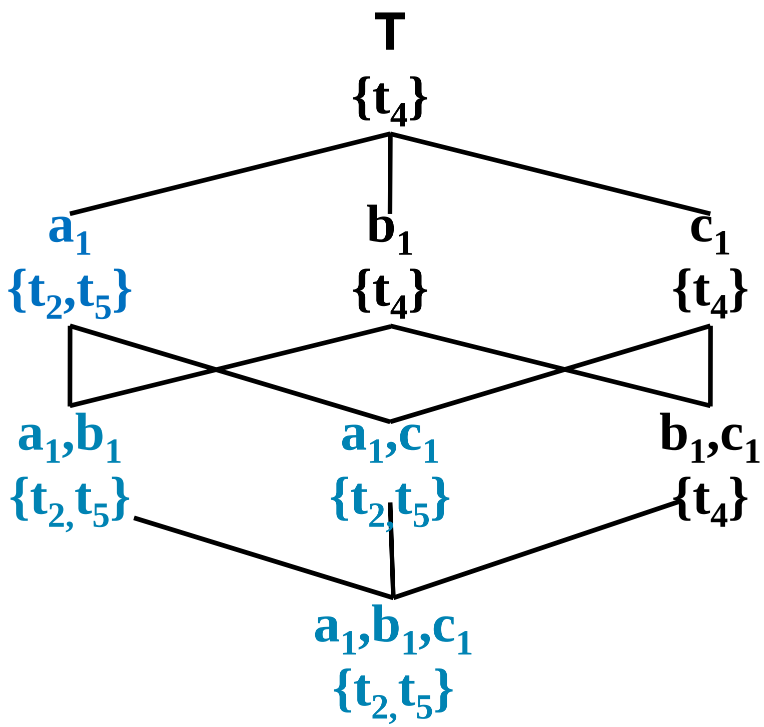
TopDown stores a tuple for its maximal skyline constraints only.

TopDown

Skyline Constraints

Constraints whose contextual skylines include *t*.

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15

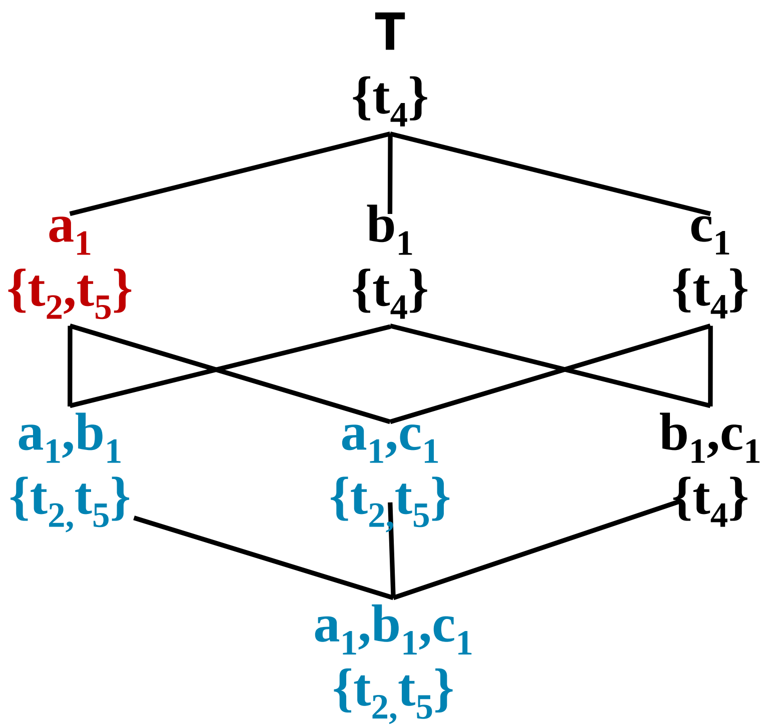


TopDown

Maximal Skyline Constraints

Constraints not subsumed by any other skyline constraints of *t*.

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15

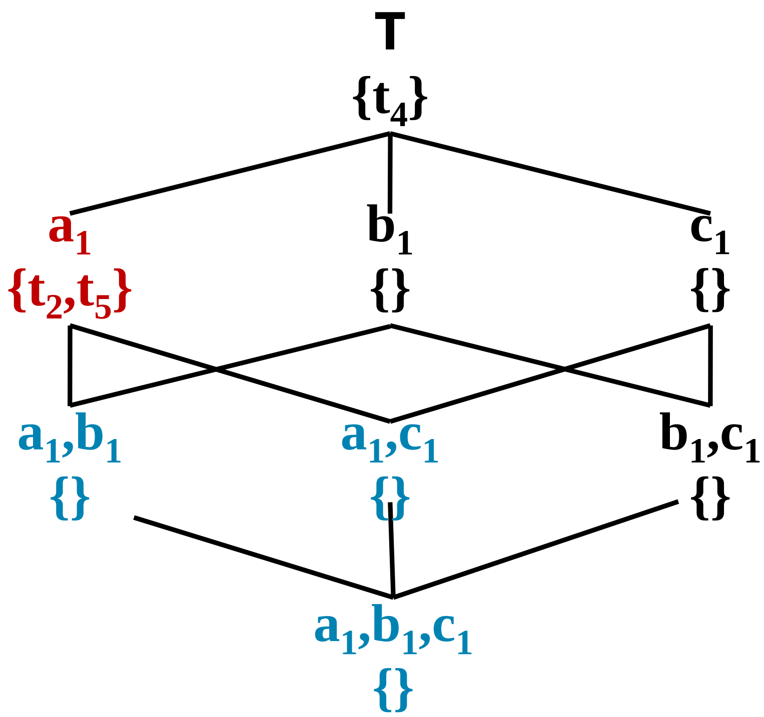


TopDown

Maximal Skyline Constraints

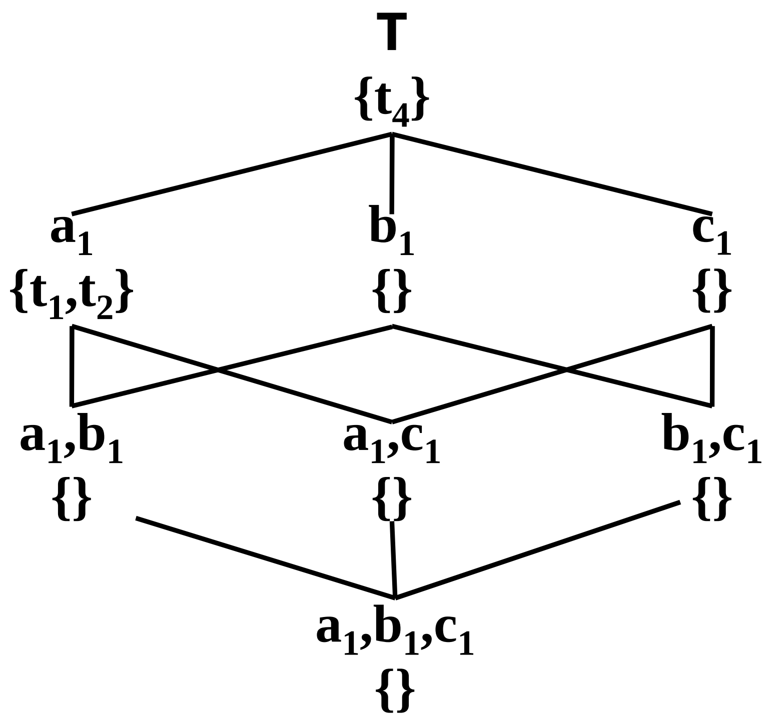
Constraints not subsumed by any other skyline constraints of *t*.

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



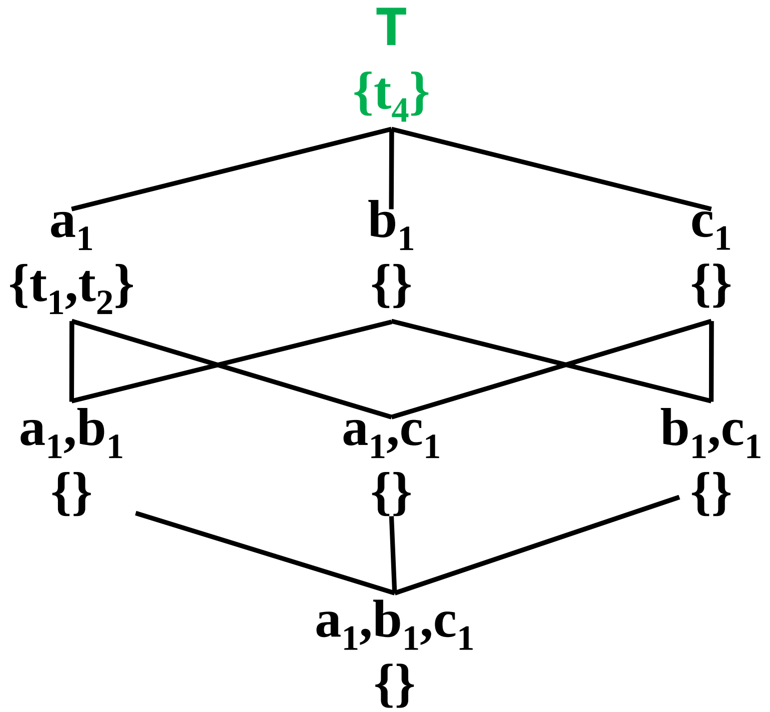
TopDown

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



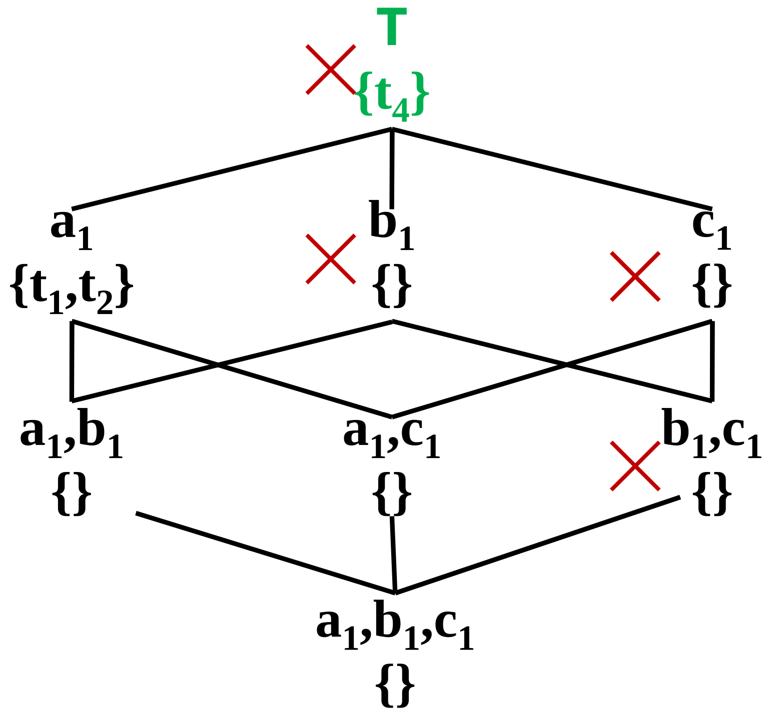
TopDown

id	*d_1*	*d_2*	*d_3*	*m_1*	*m_2*
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



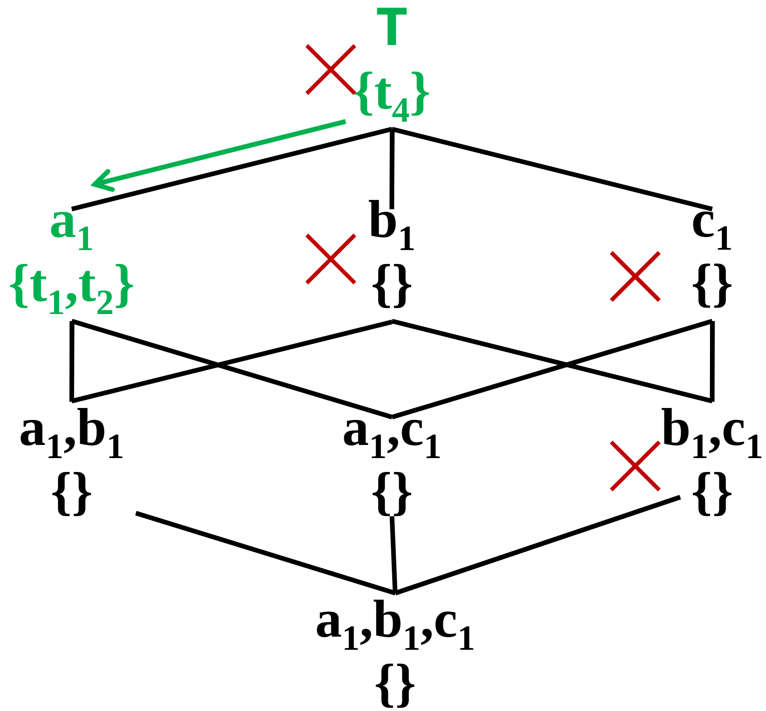
TopDown

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15



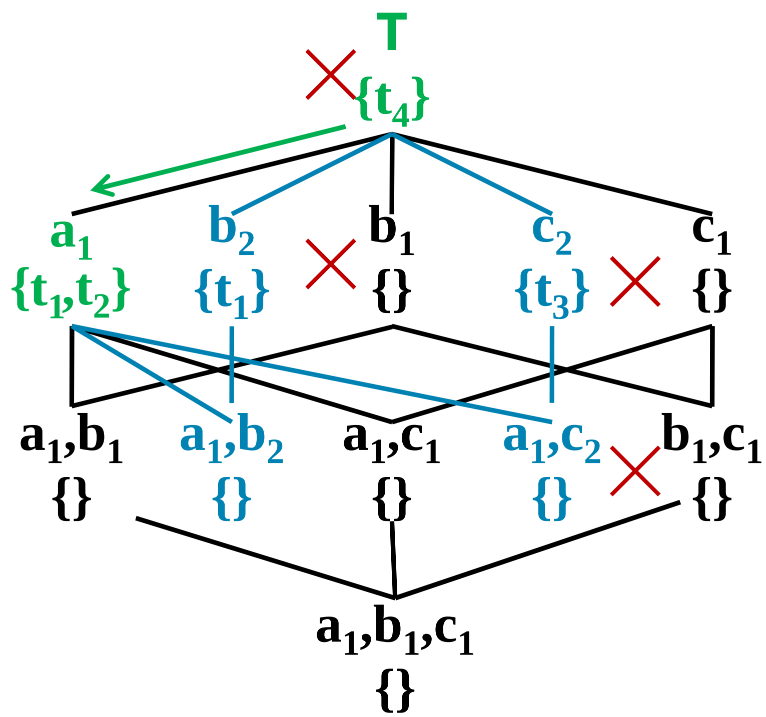
TopDown

id	d_1			m_1	m_2
t_1	a_1			10	15
t_2	a_1			15	10
t_5	a_1			11	15



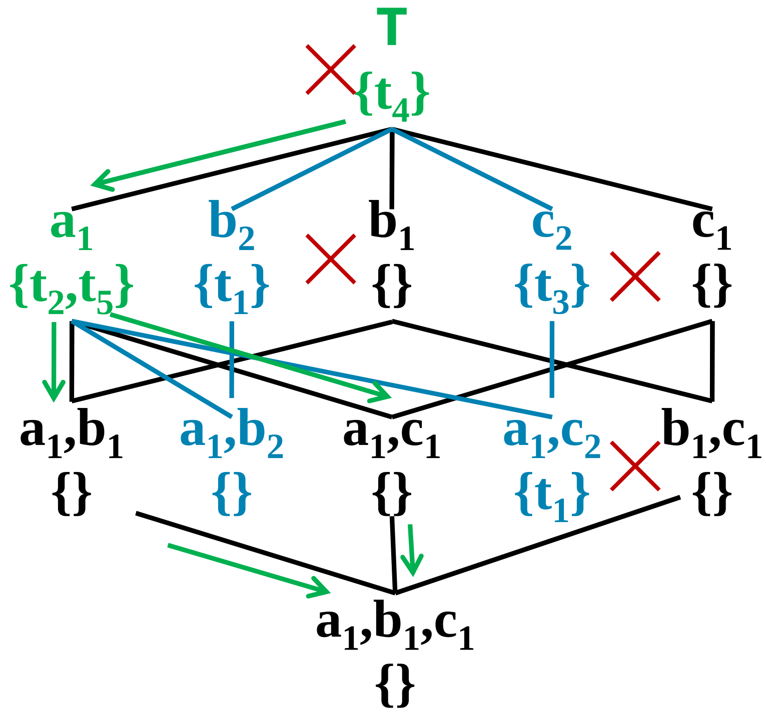
TopDown

id	d_1			m_1	m_2
t_1	a_1			10	15
t_2	a_1			15	10
t_5	a_1			11	15



TopDown

id	d_1			m_1	m_2
t_1	a_1			10	15
t_2	a_1			15	10
t_5	a_1			11	15



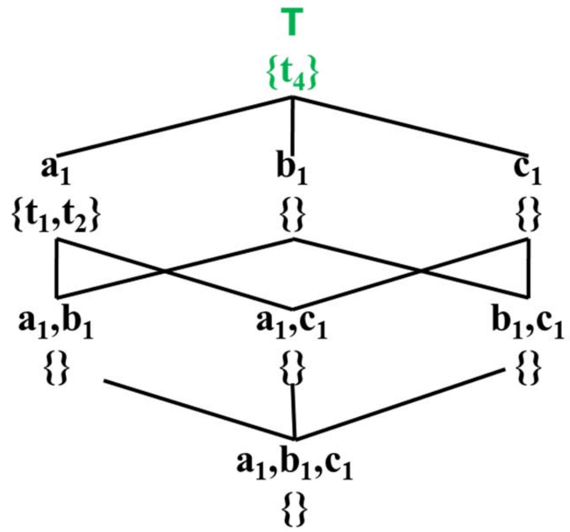
Total 3 comparisons in this case

STopDown and SBottomUp

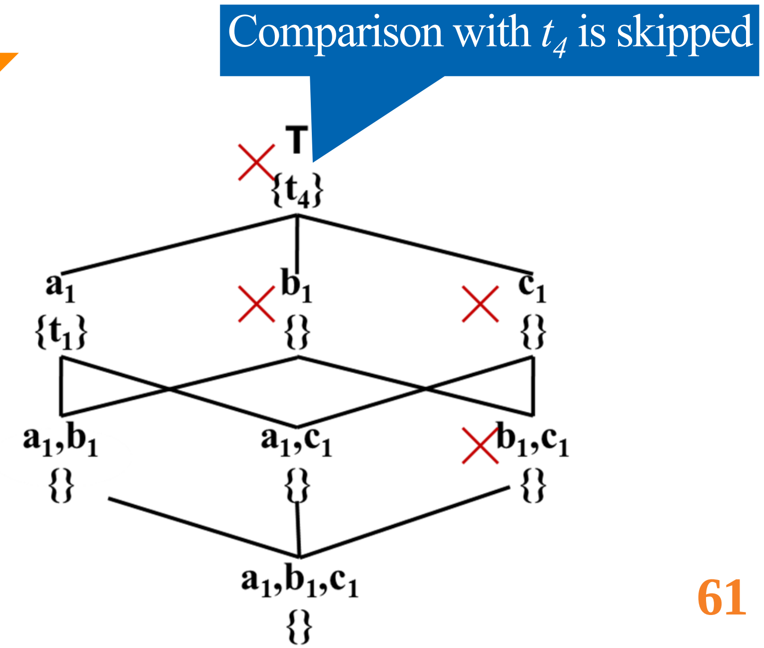
- Con of BottomUp and TopDown
  - Need to compute over every measure subspace separately
    - STopDown and SBottomUp share computation across different subspaces

STopDown

id	d_1	d_2	d_3	m_1	m_2
t_1	a_1	b_2	c_2	10	15
t_2	a_1	b_1	c_1	15	10
t_3	a_2	b_1	c_2	17	17
t_4	a_2	b_1	c_1	20	20
t_5	a_1	b_1	c_1	11	15

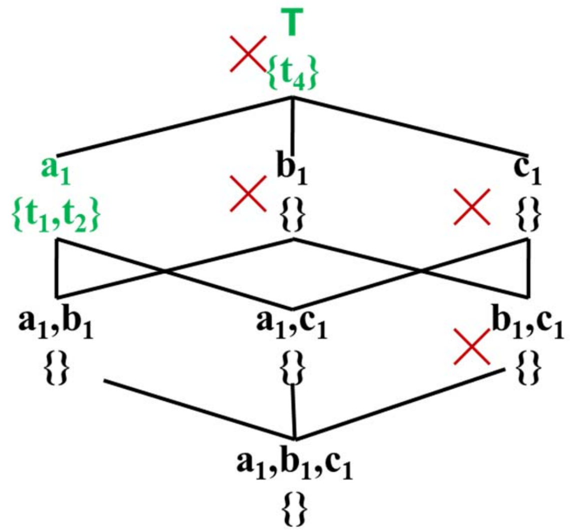


id	d_1	d_2	d_3		m_2
t_1	a_1	b_2	c_2		15
t_2	a_1	b_1	c_1		10
t_3	a_2	b_1	c_2		17
t_4	a_2	b_1	c_1		20
t_5	a_1	b_1	c_1		15

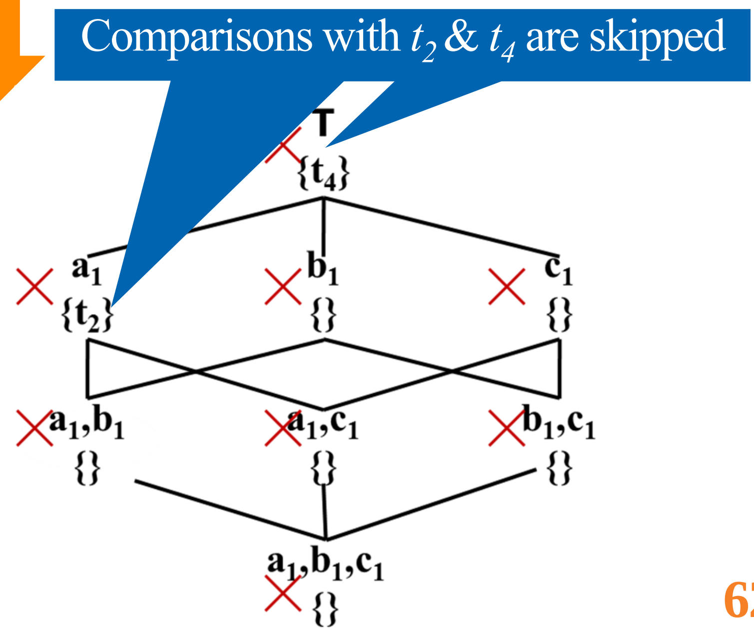


STopDown

id	d_1			m_1	m_2
t_1	a_1			10	15
t_2	a_1			15	10
t_5	a_1			11	15



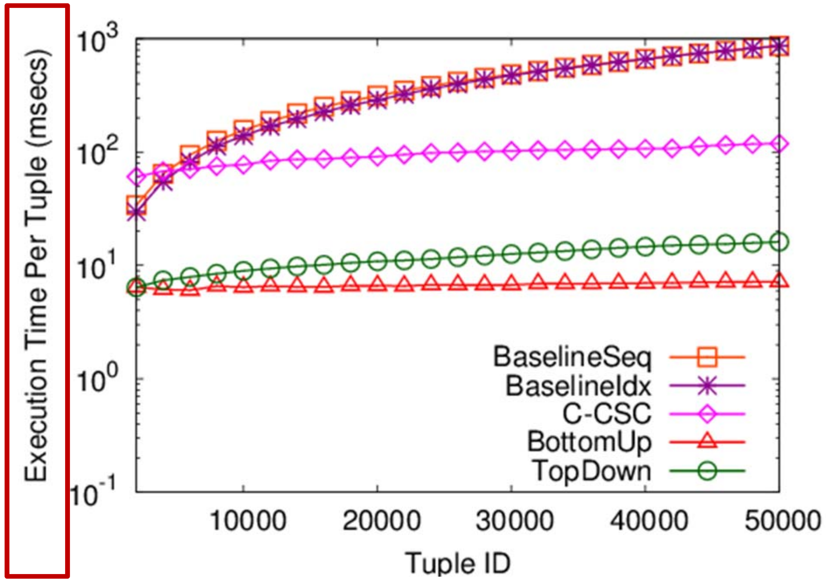
id	d_1	d_2	d_3	m_1	
t_1	a_1	b_2	c_2	10	
t_2	a_1	b_1	c_1	15	
t_3	a_2	b_1	c_2	17	
t_4	a_2	b_1	c_1	20	
t_5	a_1	b_1	c_1	11	



Experiment Setup

- NBA Dataset
  - 317,371 tuples of NBA box scores from 1991-2004 seasons
  - 8 dimension attributes
  - 7 measure attributes
- Weather Dataset
  - 7.8 million tuples of weather forecast from different locations of six countries & regions of UK
  - 7 dimension attributes
  - 7 measure attributes

Memory-Based Implementation

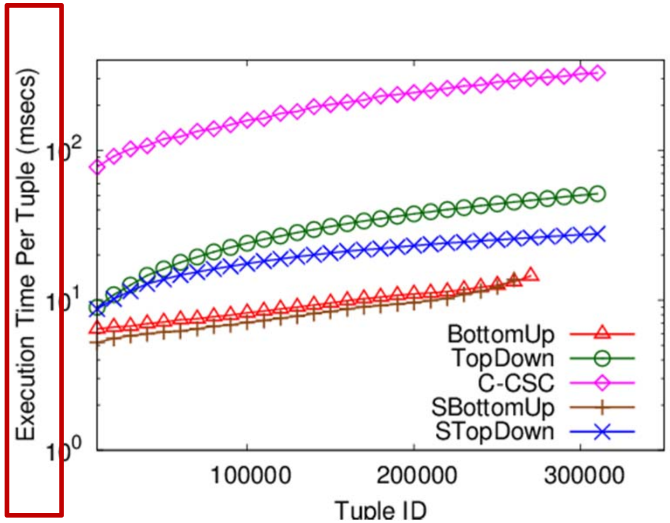


NBA Dataset

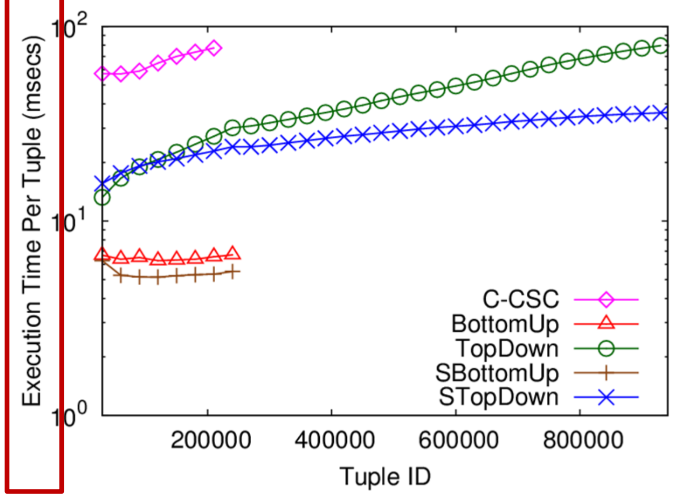
❑Maintaining CSC for each constraint causes overhead (Xia et al. SIGMOD 2006)

- Not benefitted by constraint pruning

Memory-Based Implementation



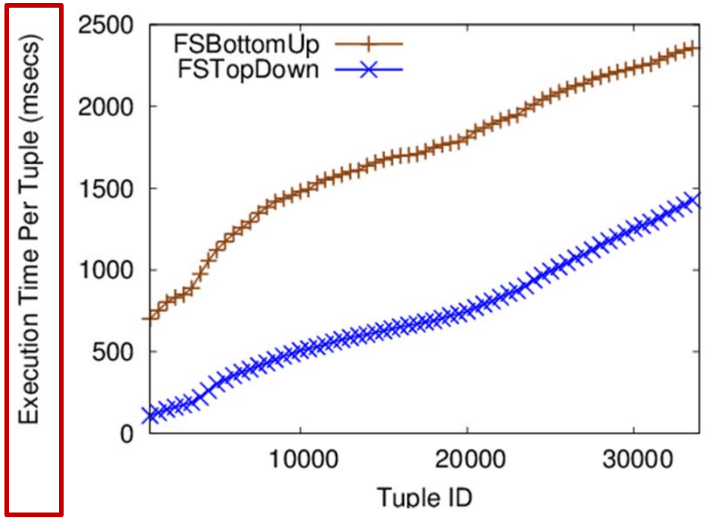
NBA Dataset



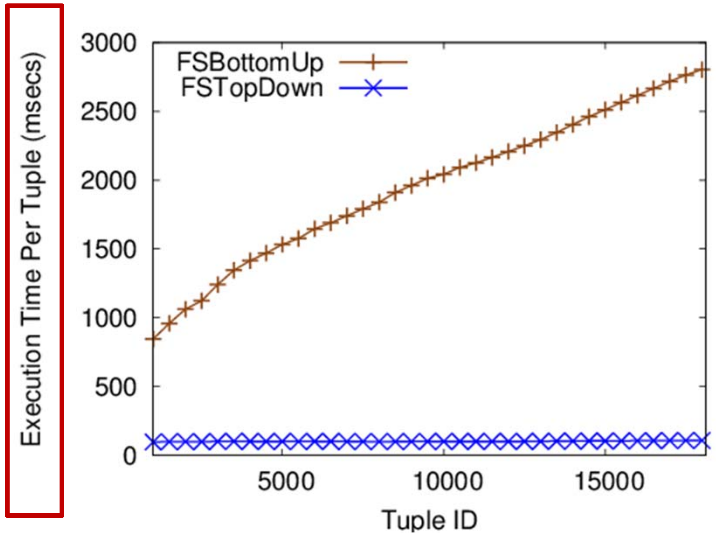
Weather Dataset

- BottomUp/SBottomUp exhausted available JVM heap
  - memory overflow
- TopDown/STopDown was outperformed by BottomUp/SBottomUp
  - Updating maximal skyline constraints causes overhead

File-Based Implementation



NBA Dataset



Weather Dataset

- Each storage of (C,M) is a binary file
- While traversing, file-read operation occurs if storage is non-empty: FSTopDown encounters many empty storage
- For updating storage, file-write operation occurs: FSTopDown stores fewer tuples
- I/O-cost dominates in-memory computation

Conclusion

Tuples capturing real world events appended to table

id	player	day	month	season	team	opp_team	pts	ast	reb
t_1	Bogues	11	Feb.	1991-92	Hornets	Hawks	4	12	5
t_2	Seikaly	13	Feb.	1991-92	Heat	Hawks	24	5	15
t_3	Sherman	7	Dec.	1993-94	Celtics	Nets	13	13	5
t_4	Wesley	4	Feb.	1994-95	Celtics	Nets	2	5	2
t_5	Wesley	5	Feb.	1994-95	Celtics	Timberwolves	3	5	3
t_6	Strictland	3	Jan.	1995-96	Blazers	Celtics	27	18	8
t_7	Wesley	25	Feb.	1995-96	Celtics	Nets	12	13	5

Find constraint-measure pair (C,M) such that t is in the contextual skyline.

Constraint	Measure
month=*Feb*	pts, ast, rb
opp_team=*Nets*	ast, rb
team=*Celtics* ∧ opp_team=*Nets*	ast, rb
…	…

Template

Wesley had 12 points, 13 assists and 5 rebounds on February 25, 1996 to become the first player with a 12/13/5 (points/assists/rebounds) in February.

- ✓ Novel problem of discovering prominent situational facts
- ✓ Presented Efficient algorithms
- ✓ Adopted prominence measure to rank

Ranking Facts

$$\boxed{\text{Prominence of Fact} = \frac{\text{All tuples}}{\text{Skyline tuple in same context}}}$$

Ranking Facts

id			**month**				**pts**	**ast**	**rb**
t_1			Feb.				4	12	5
t_2			Feb.				24	5	15
t_4			Feb.				2	5	2
t_5			Feb.				3	5	3
t_7			Feb.				12	13	5

- (month=*Feb*,{points,assists,rebounds})=>5/2

Ranking Facts

id					team	opp_team		ast	rb
t_3					Celtics	Nets		13	5
t_4					Celtics	Nets		5	2
t_7					Celtics	Nets		13	5

- $(\text{team}=\textit{Celtics}\wedge\text{opp\_team}=\textit{Nets},\{\text{assists},\text{rebounds}\})\Rightarrow 3/2$

Discovered Facts

- Lamar Odom had 30 points, 19 rebounds and 11 assists on March 6, 2004. No one before had a better or equal performance in NBA history.
- Allen Iverson had 38 points and 16 assists on April 14, 2004 to become the first player with a 38/16 (points/assists) game in the 2004-2005 season.
- Damon Stoudamire scored 54 points on January 14, 2005. It is the highest score in history made by any Trail Blazers.

Future Work

- Narrating facts in natural language text
- Demo under submission