

Modélisation, conception et exploitation de données

Temporalisation des relations

Quelles sont les bonnes façons d'inclure des attributs temporels dans une relation?

TEM_02
100b
2023-09-21

Christina.Khnaisser@USherbrooke.ca
Luc.Lavoie@USherbrooke.ca

© 2018-2021, Μητρίς (<http://info.usherbrooke.ca/lavoie>)
CC BY-NC-SA 4.0 (<https://creativecommons.org/licenses/by-nc-sa/4.0/>)

1

Plan

- Quelques acronymes
- Mise en contexte
- Problématique par l'exemple
- Problématique par l'analyse
- Les temporalités par l'exemple
- Les modèles temporels de base de données
 - État de l'art en bref
 - BCDM: *Bitemporal Conceptual Data Model*
 - DDLM: *Date, Darwen and Lorentzos Model*
 - AV: *Asserted Versioning*

2023-09-21 MCHD_TEM_02 — Temporalisation des relations (100b) © 2018-2023, Μητρίς — CC BY-NC-SA 4.0
Département d'informatique, l'École des sciences, Université de Sherbrooke, Québec

2

2

Quelques acronymes

- BDC: base de **données courantes**
 - aussi couramment appelée:
 - base de données d'exploitation,
 - base de données transactionnelle.
- BDT: base de **données temporalisées**
 - aussi couramment appelée:
 - base de données journalisées,
 - base de données temporelles.
- BDH: base de **données historicisées**
 - aussi couramment appelée:
 - base de données historiques.

Mise en contexte

- BDC et BDT
- Questions
- Et encore des questions



Quelle différence entre BDC et BDT ?

- « Two decades of research into temporal databases have unequivocally shown that a time-varying table, containing certain kinds of DATE columns, is a completely different animal than its cousin, the table without such columns. »

(Snodgrass, 2000)

- Par ailleurs, en 2004, Date, Darwen et Lorentzos vont faire la démonstration que toutes les opérations relatives à la temporalité des données s'expriment naturellement au sein de la théorie relationnelle... sans qu'il soit nécessaire d'y modifier quoi que ce soit.*

2023-09-21

MCED_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, MFrq — CC BY-NC-SA 4.0
Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

5

5

Qu'est-ce qu'une BDT ?

- A database that maintains past, present, and future data is called a temporal database.

(Tansel, 1993)

- A temporal database is one that supports some aspect of time

(Snodgrass, 1995)

- A temporal database can be thought of, very loosely, as a database that contains historical data instead of or in addition to current data.

(DDL, 2002)

2023-09-21

MCED_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, MFrq — CC BY-NC-SA 4.0
Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

6

6

Qu'est-ce qu'une BDC ?

- Une base de données qui ne contient que des données courantes (ou réputées telles).
- Il s'agit donc d'une projection d'une base de données historicisée sur l'instant présent.

(Lavoie, 2014)

Quelques questions connexes

- Qu'est-ce que le temps?
- Qu'est qu'un « évènement »?
- Peut-on récrire l'histoire?
- Que signifie « maintenant »?
- Qu'est-ce que le futur?
- Qu'est-ce que le passé?

Autres questions

- Comment définir
 - moment, période et durée?
- Quelles sont les différences entre les temps
 - calendaire, astronomique et atomique?
- Le temps est-il absolu ou relatif?
 - S'il est relatif, à quoi l'est-il?
- Y a-t-il une différence entre
 - une « information » (date de naissance) et une « temporalisation » (période de validité, temps de transaction)?
 - une information temporalisée et une qui ne l'est pas?
- Existe-t-il des informations (vraiment) immuables?
- Que signifie l'absence d'estampille temporelle dans une information?

2023-09-21

MCED_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, Mftrg — CC BY-NC-SA 4.0
 Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

9

9

Problématiques par l'exemple

- Modélisation
- Requêtes
- Modifications
- Évolution

La plus grande partie du matériel de cette section provient du cours INF 3302 SS 2018 développé par le Prof. Manthey

Voir

- [Manthey]
- [INF 3302]

2023-09-21

MCED_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, Mftrg — CC BY-NC-SA 4.0
 Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

10

10

Histoire de la présidence américaine

- Modélisation
- Requêtes
- Modifications

11

Modélisation

Quelle est la
sémantique (le sens)
des colonnes From,
To?

À quoi réfèrent les
colonnes?

Quel est leur rôle?

Presidency	President	Birthday	From	To	Term
1	George Washington	22.2.1732	30.4.1789	4.3.1793	1
...					
22	Grover Cleveland	18.3.1837	4.3.1885	4.3.1889	25
23	Benjamin Harrison	20.8.1833	4.3.1889	4.3.1893	26
24	Grover Cleveland	18.3.1837	4.3.1893	4.3.1897	27
...					
44	Barack Obama	4.8.1961	20.1.2009	20.1.2013	56
44	Barack Obama	4.8.1961	20.1.2013	20.1.2017	57
45	Donald Trump	14.6.1946	20.1.2017		58

12

Requêtes

- Qui est le président actuel ?
- Qui était le président le 4 juillet 1890 ?
- Qui sera le président le 4 juillet 2019 ?

Presidency	President	Birthday	From	To	Term
1	George Washington	22.2.1732	30.4.1789	4.3.1793	1
...					
22	Grover Cleveland	18.3.1837	4.3.1885	4.3.1889	25
23	Benjamin Harrison	20.8.1833	4.3.1889	4.3.1893	26
24	Grover Cleveland	18.3.1837	4.3.1893	4.3.1897	27
...					
44	Barack Obama	4.8.1961	20.1.2009	20.1.2013	56
44	Barack Obama	4.8.1961	20.1.2013	20.1.2017	57
45	Donald Trump	14.6.1946	20.1.2017		58

2023-09-21
 MGED_TEM_02 — Temporalisation des relations (v10b) © 2018-2023, MING — CC BY-NC-SA 4.0
 Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

13

13

Requêtes

- Quelle période est couverte par la 43^e présidence?
- Combien de temps a duré la 43^e présidence?
- Pendant combien d'années Cleveland a-t-il été président?

Presidency	President	Birthday	From	To	Term
1	George Washington	22.2.1732	30.4.1789	4.3.1793	1
...					
22	Grover Cleveland	18.3.1837	4.3.1885	4.3.1889	25
23	Benjamin Harrison	20.8.1833	4.3.1889	4.3.1893	26
24	Grover Cleveland	18.3.1837	4.3.1893	4.3.1897	27
...					
43	George W. Bush	6.7.1946	20.1.2001	20.1.2005	54
43	George W. Bush	6.7.1946	20.1.2005	20.1.2009	55
...					

2023-09-21
 MGED_TEM_02 — Temporalisation des relations (v10b) © 2018-2023, MING — CC BY-NC-SA 4.0
 Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

14

14

Modifications

- Les faits consignés dans ce tableau sont connus pour être historiquement corrects, et ne seront « jamais » réexaminés.
- Les seules modifications à prévoir sont l'ajout de nouvelles lignes et la saisie de nouvelles dates manquantes.
- ... Vraiment?

Presidency	President	Birthday	From	To	Term
1	George Washington	22.2.1732	30.4.1789	4.3.1793	1
...					
22	Grover Cleveland	18.3.1837	4.3.1885	4.3.1889	25
23	Benjamin Harrison	20.8.1833	4.3.1889	4.3.1893	26
24	Grover Cleveland	18.3.1837	4.3.1893	4.3.1897	27
...					
44	Barack Obama	4.8.1961	20.1.2009	20.1.2013	56
44	Barack Obama	4.8.1961	20.1.2013	20.1.2017	57
45	Donald Trump	14.6.1946	20.1.2017		58

Évaluation de cours dans une université

- Évolution des données
- Modification des données

Évolution (1)

Les données du
1^{er} avril 2011...

Qu'y a-t-il d'étrange?

Comment l'expliquer?

Student	Class	Signed_up	Dropped	Grade	Exam Date
John	1203	11.11.2010		1,3	13.2.2011
Jack	1203	19.11.2010	2.1.2011		
Tim	1203	21.11.2010		3,0	18.3.2011
Pete	1203	27.11.2010	3.2.2011	5,0	18.3.2011
John	2201	11.11.2010		1,7	19.2.2011
Jack	2201		2.1.2011		
Tim	3203	2.12.2010		3,7	1.4.2011

17

Évolution (2)

Les données au
4 avril 2011...

Student	Class	Signed_up	Dropped	Grade	Exam Date
John	1203	11.11.2010		1,3	13.2.2011
Jack	1203	19.11.2010	2.1.2011		
Tim	1203	21.11.2010		3,0	18.3.2011
Pete	1203	27.11.2010	3.2.2011	5,0	18.3.2011
John	2201	11.11.2010		1,7	19.2.2011
Jack	2201		2.1.2011		
Tim	3203	2.12.2010		3,7	1.4.2011

18

Évolution (3)

Les données du 6
avril 2011...

Student	Class	Signed_up	Dropped	Grade	Exam Date
John	1203	11.11.2010		1,3	13.2.2011
Jack	1203	19.11.2010	2.1.2011		
Tim	1203	21.11.2010		3,0	18.3.2011
Pete	1203	27.11.2010	3.2.2011	5,0	18.3.2011
John	2201	11.11.2010		1,7	19.2.2011
Jack	2201		2.1.2011		
Tim	3203	2.12.2010		3,7	1.4.2011

Évolution (4)

Une solution?

Student	Class	Signed_up	Dropped	Grade	Exam Date	From	To
John	1203	11.11.2010				11.11.2010	14.2.2011
John	1203	11.11.2010		1,3	13.2.2011	14.2.2011	
Jack	1203	19.11.2010				19.11.2010	2.1.2011
Jack	1203	19.11.2010	2.1.2011			2.1.2011	
Tim	1203	21.11.2010				21.11.2010	20.3.2011
Tim	1203	21.11.2010		3,0	18.3.2011	20.3.2011	8.4.2011
Tim	1203	21.11.2010		2,7	18.3.2011	8.4.2011	
Pete	1203	27.11.2010				27.11.2010	3.2.2011
Pete	1203	27.11.2010	3.2.2011			3.2.2011	21.3.2011
Pete	1203	27.11.2010	3.2.2011	5,0	18.3.2011	21.3.2011	6.4.2011
Pete	1203	27.11.2010	3.2.2011			6.4.2011	
John	2201	11.11.2010				11.11.2010	21.2.2011
John	2201	11.11.2010		1,7	19.2.2011	21.2.2011	
Jack	2201		2.1.2011			2.1.2010	7.4.2011
Tim	3203	2.12.2010		3,7	1.4.2010	1.4.2011	4.4.2011
Tim	3203	2.12.2010		2,7	1.4.2011	4.4.2011	

Problématiques par l'analyse

- Redondance
- Circonlocution
- Contradiction
- Non-compacité

2023-09-21

MCED_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, Mfres — CC BY-NC-SA 4.0
 Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

21

21

Redondance

- Les intervalles de deux tuples contenant les mêmes données partagent un point.

Séjour			
noPatient	noUnite	noChambre	d
P1	U1	1200	[d02:d08]
P1	U1	1200	[d04:d10]

2023-09-21

MCED_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, Mfres — CC BY-NC-SA 4.0
 Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

22

22

Circonlocution

- Les intervalles de deux tuples contenant les mêmes données se jouxtent.

Séjour			
noPatient	noUnite	noChambre	d
P1	U1	1200	[d02:d07]
P1	U1	1200	[d08:d10]

2023-09-21

MCHD_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, Mftrg — CC BY-NC-SA 4.0
Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

23

23

Contradiction

- Les intervalles de deux tuples ayant la même clé et des données différentes partagent un point.

Séjour			
noPatient	noUnite	noChambre	d
P1	U1	1200	[d02:d08]
P1	U1	1300	[d04:d10]

2023-09-21

MCHD_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, Mftrg — CC BY-NC-SA 4.0
Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

24

24

Non-compacité

- La compacité (*densness*) est une assertion qui garantit la cohérence des clés référentielles.
- Soit la clé référentielle
 $\text{Sejour}\{\text{noPatient}\} \rightarrow \text{Hospitalisation}$

Hospitalisation	
noPatient	d
P1	[d02:d20]

Séjour			
noPatient	noUnite	noChambre	d
P1	U2	2400	[d02:d23]

2023-09-21

MCEID_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, MFRG — CC BY-NC-SA 4.0
 Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

25

25

Compacité référentielle et compacité prédicative

- Nous distinguerons plus tard deux formes de compacité, la référentielle (présentée ici) et la prédicative (lorsque nous aborderons la modélisation en 6^e forme normale).
- À proprement parler, la compacité n'est généralement requise qu'au moment de l'historicisation de la base de données.

2023-09-21

MCEID_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, MFRG — CC BY-NC-SA 4.0
 Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

26

26

Les temporalités par l'exemple

- Atemporalité
- Unitemporalité
- Bitemporalité
- Unitemporalité vs Bitemporalité
- Multitemporalité et agents

2023-09-21

MCED_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, M@rs — CC BY-NC-SA 4.0
 Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

27

27

Exemple Hospitalisation Modèle de base

- Admission à l'hôpital et assignation d'une unité (de soin) et d'une chambre.
- Une modélisation *atemporelle* simple

Patient			
noPatient	nom	ville	naissance

Séjour		
noPatient	unité	chambre

2023-09-21

MCED_TEM_02 — Temporalisation des relations (v100b) © 2018-2023, M@rs — CC BY-NC-SA 4.0
 Département d'informatique, Faculté des sciences, Université de Sherbrooke, Québec

28

28

Exemple Hospitalisation Modèle de base - insuffisance

Le patient Tremblay (12345), né le 3 octobre 1960 et domicilié à Sherbrooke, est hospitalisé en cardiologie dans la chambre 3210.

Patient			
noPatient	nom	ville	naissance
12345	Tremblay	Sherbrooke	1960-10-03

Séjour		
noPatient	unité	Chambre
12345	Cardiologie	3210

29

Exemple Hospitalisation Modèle de base - insuffisance

Le patient Tremblay (12345), né le 3 octobre 1960 et domicilié à Sherbrooke, est hospitalisé en cardiologie dans la chambre 3210.

Patient			
noPatient	nom	ville	naissance
12345	Tremblay	Sherbrooke	1960-10-03

Séjour		
noPatient	unité	Chambre
12345	Cardiologie	3210

Le patient Tremblay obtient son congé et quitte l'hôpital.

Patient			
noPatient	nom	ville	naissance
12345	Tremblay	Sherbrooke	1960-10-03

Séjour		
noPatient	unité	Chambre

Quand?

A-t-il déjà
été
hospitalisé?

30

Exemple Hospitalisation Modèle unitemporel

Le patient Tremblay, né le 3 octobre 1960 et domicilié à Sherbrooke, est hospitalisé en cardiologie dans la chambre 3210 *pour la période commençant le d02 et se terminant le d05*.

Patient				
noPatient	nom	ville	naissance	dValidite
12345	Tremblay	Sherbrooke	1960-10-03	[d02:??]

Séjour			
noPatient	unité	Chambre	dValidite
12345	Cardiologie	3210	[d02:d05]

Congé: Le patient Tremblay obtient son congé le **d04** (non le **d05** comme prévu) et quitte l'hôpital.

Patient				
noPatient	nom	ville	naissance	dValidite
12345	Tremblay	Sherbrooke	1960-10-03	[d02:??]

Séjour			
noPatient	unité	chambre	dValidite
12345	Cardiologie	3210	[d02:d04]

31

Exemple Hospitalisation Modèle bitemporel

La BD enregistre le d02 «Le patient Tremblay, né le 3 octobre 1960 et domicilié à Sherbrooke, est hospitalisé en cardiologie dans la chambre 3210 *pour la période commençant le d02 et se terminant le d05*.»

Patient					
noPatient	nom	ville	naissance	dValidite	dTransaction
12345	Tremblay	Sherbrooke	1960-10-03	[d02:??]	[d02:??]

Séjour				
noPatient	unité	Chambre	dValidite	dTransaction
12345	Cardiologie	3210	[d02:d05]	[d02:??]

La BD enregistre le d02 «Le patient Tremblay obtient son congé le **d04** (non le **d05** comme prévu) et quitte l'hôpital.»

Séjour				
noPatient	unité	chambre	dValidite	dTransaction
12345	Cardiologie	3210	[d02:d05]	[d02:d03]
12345	Cardiologie	3210	[d02:d04]	[d04:??]

32

Exemple Hospitalisation Modèle unitemporel VS bitemporel

- Temps de validité: intervalle de temps durant lequel la proposition associée est considérée vrai

Séjour			dValidite
noPatient	unité	Chambre	
12345	Cardiologie	3210	

- Temps de transaction: intervalle de temps durant lequel une proposition est présente dans la base de données

Séjour				dTransaction
noPatient	unité	Chambre	dValidite	
12345	Cardiologie	3210	[d02:d05]	

Exemple Hospitalisation Modèle unitemporel VS bitemporel

Séjour			dValidite	dTransaction
noPatient	unité	Chambre		
12345	Cardiologie	3210	[d02:d05]	[d02:d03]

- Au temps d2:
 - Le patient 12345 occupe la chambre 3210 au service de Cardiologie (au temps d2);
 - Le patient 12345 occupera la chambre 3210 au service de Cardiologie au temps d3;
 - Le patient 12345 occupera la chambre 3210 au service de Cardiologie au temps d4;
 - Le patient 12345 occupera la chambre 3210 au service de Cardiologie au temps d5.
- Au temps d3:
 - Le patient 12345 a occupé la chambre 3210 au service de Cardiologie au temps d2;
 - Le patient 12345 occupe la chambre 3210 au service de Cardiologie (au temps d3);
 - Le patient 12345 occupera la chambre 3210 au service de Cardiologie au temps d4;
 - Le patient 12345 occupera la chambre 3210 au service de Cardiologie au temps d5.

où chaque proposition est le «point de vue» du modèle instancié par la base de données.

Les colles du prof

- Écrire les assertions requises pour empêcher la modification incorrecte d'une BD en regard des quatre problématiques temporelles. La BD est formée des trois relations Patient, Hospitalisation et Séjour.
 - Utiliser
 - un langage calqué sur l'algèbre relationnelle (comme Tutorial D ou Discipulus);
 - le langage SQL (variante ISO ou PostgreSQL).
 - Faire l'exercice pour les trois variantes
 - unitemporelles de validation;
 - unitemporelle de transaction;
 - bitemporelle.

Références

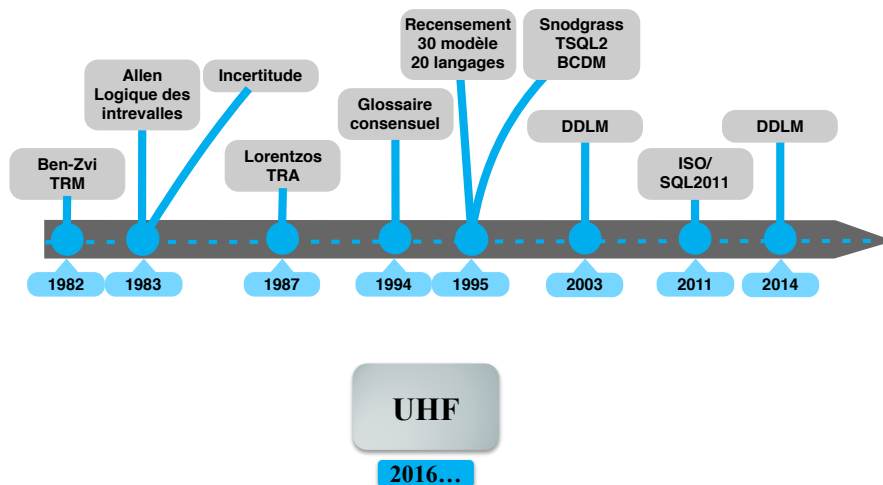
- Théorie temporelle
 - C.J. Date, H. Darwen, N.A. Lorentzos. 2014.
Time and the relational theory.
 Morgan Kaufman.

Les modèles temporels de bases de données

- État de l'art en bref
- Modèles:
 - Temporalité générale
 - BCDM: Bitemporal Conceptual Data Model
 - DDLM: Date, Darwen and Lorentzos Model
 - AV: Asserted Versioning

37

Les modèles temporels de bases de données État de l'art en bref



38

Temporalité générale

- [Manthey]
https://pages.iai.uni-bonn.de/manthey_rainer/
- [INF 3302]
https://pages.iai.uni-bonn.de/manthey_rainer/TIS2018/

BCDM

○ *Bitemporal Conceptual Data Model*

- proposé par
JENSEN, C.S., SOO, M.D., AND SNODGRASS, R.T. 1993.
Unifying Temporal Data Models via a Conceptual Model.
Information Systems 19, 513–547.
- Développé par
SNODGRASS, R.T. 2000.
Developing time-oriented database applications in SQL.
Morgan Kaufmann Publishers, San Francisco, Calif.
<http://www.cs.arizona.edu/people/rts/tdbbook.pdf>
- bien présenté par
MANTHEY, R. 2014.
Temporal Information Systems
<http://www.iai.uni-bonn.de/III/lehre/vorlesungen/TemporalIS/SS14/>

DDL

Temporal Relational Model

- proposé par
LORENTZOS, N.A. AND JOHNSON, R.G. 1988.
Extending relational algebra to manipulate temporal data. Information Systems 13, 3, 289–96.
- développé par
- DATE, C.J., DARWEN, H., AND LORENTZOS, N.A. 2003.
Temporal data and the relational model: a detailed investigation into the application of interval and relation theory to the problem of temporal database management. Morgan Kaufmann Publishers, San Diego, CA.
- DATE, C.J., DARWEN, H., AND LORENTZOS, N.A. 2014.
Time and relational theory: temporal databases in the relational model and SQL.
Morgan Kaufman/Elsevier, Amsterdam ; Boston.

AV

Asserted Versioning

- JOHNSTON, T. AND WEIS, R. 2010.
Managing time in relational databases: how to design, update and query temporal data.
Morgan Kaufmann/Elsevier, Amsterdam ; Boston.

