



HAL
open science

UN MODELE DE GRAPHE SPATIO-TEMPOREL POUR REPRESENTER L'EVOLUTION D'ENTITES GEOGRAPHIQUES

Géraldine del Mondo

► **To cite this version:**

Géraldine del Mondo. UN MODELE DE GRAPHE SPATIO-TEMPOREL POUR REPRESENTER L'EVOLUTION D'ENTITES GEOGRAPHIQUES. Autre [cs.OH]. Université de Bretagne occidentale - Brest, 2011. Français. NNT: . tel-00651682

HAL Id: tel-00651682

<https://theses.hal.science/tel-00651682>

Submitted on 14 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THESE / UNIVERSITE DE BREST
sous le sceau de l'Université européenne de Bretagne

pour obtenir le titre de
DOCTEUR DE L'UNIVERSITE DE BREST
Mention : Informatique - Géomatique
Ecole Doctorale des Sciences de la Mer

présentée par
Géraldine Del Mondo

Préparée à l'Institut de Recherche de l'Ecole Navale
(EA 3634)

Un modèle de graphe spatio-temporel pour représenter l'évolution d'entités géographiques



Thèse soutenue le 14 octobre 2011
devant le jury composé de :

Jacques Tisseau
Professeur à l'Ecole Nationale des Ingénieurs de Brest / *président*

Jérôme Gensei
Professeur à l'université Pierre Mendès France,
Laboratoire d'Informatique de Grenoble / *rapporteur*

Florence Le Ber
Chargée de recherche au Laboratoire d'Hydrologie et de
Géochimie de Strasbourg / *rapporteur*

Roland Billen
Professeur à l'Université de Liège / *examineur*

Gerard Ligozat
Professeur à l'université Paris-Sud / *invité*

John Stell
Enseignant chercheur à l'université de Leeds / *invité*

Christophe Claramunt
Professeur à l'Institut de Recherche de l'Ecole Navale / *directeur de thèse*

Rémy Thibaud
Enseignant chercheur à l'Institut de Recherche de l'Ecole Navale / *co-directeur de thèse*

THÈSE

présentée à

L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE

Mention *Informatique - Géomatique*

par

Géraldine Del Mondo

Un modèle de graphe spatio-temporel pour
représenter l'évolution d'entités géographiques

soutenue le 14 octobre 2011 devant la commission d'examen :

Composition du Jury

Rapporteurs : Pr. Jérôme Gensel, Université Pierre Mendès France (LIG), Grenoble
Dr. Florence Le Ber, LHYGES Strasbourg

Examineurs : Pr. Roland Billen, ULG Liège
Pr. Jacques Tisseau, ENIB Brest

Invités : Pr. Gérard Ligozat, Université Paris-sud
Dr. John Stell, Université de Leeds

Directeur : Pr. Christophe Claramunt, IRENav Brest

Co-Directeur : Dr. Rémy Thibaud, IRENav Brest

Remerciements

Des entités qui évoluent dans l'espace et le temps et qui interagissent, une vie en somme! Modéliser une vie c'est un peu long, mais il y a un certains nombres d'entités que je voudrais particulièrement remercier, car elles sont partie prenante du processus qui m'a conduit à obtenir ce doctorat aujourd'hui. Tout a bien évidemment commencé avec un type d'entités particulier, le type *famille*. Je remercie toutes les entités de ce type pour leur soutien, et en particulier mes parents qui m'ont fait confiance alors qu'après une scolarité sans faille, mon cerveau s'est mis en veille quelques années¹... On pourrait spécifier une fonction pour identifier cet interlude dans l'évolution de vie que je modélise ici, mais c'est sans doute trop sémantique, "*context dependant*" pour que mon modèle puisse y analyser quelque chose.

Bref, passons à l'évènement responsable de mon attrait pour la théorie des graphes, \forall les entités présentes dans ce paragraphe, je les remercie très sincèrement. Cet évènement se situe à $t = 2^{eme}$ année de DEUG lors du cours du *Pr. Jean-Xavier Rampon*, c'est dans cet amphi à l'université de Nantes que je me suis dit que les graphes étaient quelque chose qu'il devait être fabuleux de pouvoir manipuler au quotidien. Mes premiers pas dans la recherche en bio-informatique avec le *Pr. Irena Rusu* et le *Dr. Damien Eveillard* m'ont définitivement démontré mon intérêt pour la recherche scientifique. Ces années à la FAC, c'est aussi tous ces amis qui m'ont soutenue et me soutiennent encore aujourd'hui : *Jimmy, Alain, Sébastien* (Alias Boub..), *Bertrand*... Et bien sur les entités présentes depuis $t - 1$, et toujours bien là : *Angela, Songül, Adeline*...

Temps suivant $t = \{\text{Novembre 2008 - Septembre 2011}\}$. Trois années localisées à l'Institut de Recherche de l'Ecole Navale de Lanvéoc Poulmic, institut qui m'a permis de poser les premières briques de ma vie professionnelle dans la recherche et l'enseignement. Ici de nombreuses entités à remercier, je commencerai par le *Pr. Christophe Claramunt* et le *Dr. Rémy Thibaud* pour leurs conseils, leur sympathie et leur patience, ainsi que pour les opportunités qu'ils ont su me faire

1. J'entends d'ici les médisants qui diront qu'il est resté en veille ;-)

partager. En particulier mes collaborations avec le *Dr. John Stell* de l'université de Leeds, ainsi qu'avec le *Pr. Andrea Rodriguez* de l'université de Concepcion, chercheurs avec qui j'ai beaucoup aimé travailler et qui ont contribué, chacun dans leur domaine, à mon apprentissage scientifique.

Une entité se détache particulièrement, et s'il on passe à une granularité temporelle plus fine, elle se situe à $t =$ décembre 2008, date à laquelle j'ai ajouté des liens de filiations dans mon graphe. Merci *Thomas* d'être à mes côtés depuis bientôt 3 ans, pour tout le bonheur et le soutien que tu m'apportes au quotidien. Un normand avec une pseudo-ligero-italo-corse, moi je trouve que cela va très bien ensemble.

Soit $P = \{e, \text{entité} \mid e \in \text{IRENav} \cup \text{EN}\}$, je remercie sincèrement toutes les entités de cet ensemble² P pour leur accueil et leur sympathie durant ces années de thèse. En particulier, dans P on peut ou l'on a pu trouver entre autres³ : *Benoît*(surfman), *Christophe* (java, saucisson), *Cyril*(RP2EN), *Dieudonné*(sport), *Emmery*(cool), *Eric*(math, voile), *Joseph*(crabe), *Georges*(zenitude), *Imad*(carotène), *Jacques André*(corsica, corsica), *Jean-Marie*(collègue de bureau 1, relecteur), *Laurent*(collègue de bureau 2, pompier etc.), *Magalie*(efficacité, sourire), *Marie*(relectrice, future LateXienne), *Marie-France*(cupcake), *Marie-Laure*(noctambule), *Mathieu*(LateXien, gâteau au soja), *Meriam*(amitié sans édulcorant), *Pierre-Luc*(volleyman, tartiflette), *Sébastien*(biscuiterie, Bray dunes), *Thierry*(monsieur gentil).

Tout ce monde pour arriver à $t = 14$ octobre 2011. Je remercie ici l'ensemble des entités qui composent mon jury : mes deux rapporteurs le *Dr. Florence Le Ber* et le *Pr. Jérôme Gensel*, mes deux examinateurs le *Pr. Roland Billen* et le *Pr. Jacques Tisseau*, ainsi que mes deux invités le *Pr. Gerard Ligozat* et le *Dr. John Stell*, pour le temps et l'intérêt qu'ils ont consacré à mon manuscrit ainsi que pour leurs conseils.

$t =$ Octobre 2011. Géolocalisée à l'INSA de Rouen...

Et voilà, mais comme toute modélisation de phénomène réel, je ne m'avance pas sur la complétude des données du graphe spatio-temporel que je viens de décrire... alors mille excuses sincères à toutes les entités qui ne seraient pas apparues dans cette modélisation... tout ça c'est du discret, la vie non.

2. Comme ça je suis sûre de n'oublier personne, c'est beau les maths ;-)

3. Par ordre alphabétique *NomEntité* (un/deux attribut(s) parmi tant d'autres !)

Table des matières

Table des matières	i
Liste des figures	v
Liste des tableaux	xiii
Introduction	1
I Etat de l'art	7
I.1 Entité et identité	7
I.2 Modélisation des entités spatio-temporelles	10
I.2.1 Conception de l'espace et du temps	10
I.2.1.1 L'espace	10
I.2.1.2 Le temps	13
I.2.2 Dimension spatio-temporelle	17
I.2.2.1 Les modèles basés sur une approche discrète ou continue	19
I.2.2.2 Les modèles basés sur les concepts d'évènements et de processus	25
I.3 Les graphes comme support de modélisation spatio-temporelle	30
II Un graphe spatio-temporel pour modéliser l'évolution d'entités	35
II.1 Spécification du modèle	36
II.1.1 Relations de l'espace et du temps	38

II.1.1.1	Définition du domaine temporel	38
II.1.1.2	Relation de connexion spatiale	38
II.1.1.3	Relation spatio-temporelle	39
II.1.1.4	Relation de filiation	40
II.1.2	Le graphe spatio-temporel	42
II.1.3	Un exemple de modélisation : propagation d'une population de ronces	44
II.2	Fonctions du modèle	47
II.2.1	Fonctions basées sur les relations	48
II.2.1.1	Définition générale	48
II.2.1.2	Fonctions basées sur la relation spatiale de connexion	49
II.2.1.3	Fonctions basées sur la relation spatio-temporelle	50
II.2.1.4	Fonctions basées sur la relation de filiation	51
II.2.1.5	Application des fonctions basées sur les relations à l'exemple sur l'évolution de la population de ronces	52
II.2.2	Fonctions d'identification des processus	55
II.2.3	Fonction de jointure	56
II.2.3.1	Fonctions d'union et d'intersection	57
II.2.3.2	Fonction de jointure conditionnelle	57
II.2.3.3	Un exemple d'application de la jointure	59
II.2.4	Routes et relations	64
II.2.4.1	Définition générale et routes spatio-temporelles	64
II.2.4.2	Exemple d'utilisation des routes	68
III	Contraintes	71
III.1	Motivation	71
III.2	Description du formalisme pour la spécification des contraintes	73
III.2.1	Schéma général	74

III.2.2 Entité	75
III.2.3 Relations topologiques	77
III.2.4 Relations de filiations	77
III.2.5 Le graphe spatio-temporel	78
III.2.6 Fonctions de jointure	79
III.3 Spécification des contraintes	81
III.3.1 Les contraintes inhérentes au modèle : contraintes de domaine(<i>DC</i>)	81
III.3.2 Les contraintes dépendantes d'une application : contraintes sémantiques(<i>SC</i>)	84
III.4 Raisonner et vérifier la consistance	87
III.4.1 <i>Satisfiabilité</i> de l'ensemble de contraintes	87
III.4.2 Algorithme de consistance d'un ensemble de contraintes sur une instance de graphe	87
IV Prise en compte de la granularité spatiale	91
IV.1 Motivations	91
IV.2 La granularité spatiale dans le modèle de graphe spatio-temporel	95
IV.3 La piste des bigraphes	101
IV.3.1 Définition d'un bigraphe	102
IV.3.2 Application à la problématique spatio-temporelle : dynamique dans les bigraphes	106
Conclusion	113
Annexes	121
Annexe I : Implémentation du graphe spatio-temporel	122
Annexe II : Application aux données de dunes marines	124
Annexe III : Preuve de consistance	129
Bibliographie	131

Liste des figures

1	Contextualisation de la thèse suivant les différents domaines scientifiques abordés.	4
2	Méthodologie employée par cette recherche en vue de l'étude d'un phénomène spatio-temporel	5
3	(a) Relations topologiques du <i>RCC8</i> ; (b) Treillis des relations topologiques.	11
4	Exemples de relations topologiques entre deux entités dont le type d'empreinte spatiale est différent (surface et ligne/point) - Figure adaptée de [Kurata, 2009] : (a) spécification de l'extérieur E^- , intérieur E^o , frontière ∂E pour chaque entité; (b) cas d'une surface C et d'une ligne L , une case de la matrice est grisée s'il existe une relation topologique de connexion; (c) cas de la surface et du point P qui ne possède ni intérieur, extérieur ou frontière, on identifie uniquement la partie de la surface avec laquelle il est en relation topologique	12
5	Structures de temps envisageables selon le type de phénomène à représenter (a) temps cyclique; (b) le temps se divise en deux ("branching time") à l'instant e , par conséquent des évènements peuvent se produire en parallèle ensuite	14
6	Algèbre temporel de Allen : à l'exception des intervalles égaux, toutes les relations de X vers Y peuvent être établies du point de vue de Y (i.e. X avant Y , Y après X), d'où les 13 relations	15
7	Diagramme de Hasse pour l'ordonnancement qualitatif d'évènements ([Frank, 1994])	16
8	Le modèle snapshot selon [Armstrong, 1988]. Dans cet exemple, une entité est une parcelle de terrain possédant une certaine propriété (e.g. type de la culture qui s'y trouve)	20

9	Exemple de modèle Space-Time Composites modélisant un feu de forêt : la matrice décrit l'évolution du feu dans le temps, 0 (resp. 1) indique une zone non brûlée (resp. brûlée) au temps t considéré - Figure adaptée de [Langran et Chrisman, 1988]	21
10	Le modèle Spatio-temporal Object selon ([Worboys, 1992]). Exemple d'interprétation : l'entité S_1 peut être décomposée en deux atomes qui représentent spatialement l'entité en t_1 et t_2	22
11	Quelques relations du <i>RCC8</i> vu par [Egenhofer et Al-Taha, 1992]. Si l'on considère ces trois situations à trois temps consécutifs, l'évolution des relations spatiales entre X et Y se lit dans les matrices associées.	23
12	Quelques primitives de base du langage CDL [Hornsby et Egenhofer, 2000] : (a) entité qui n'existe pas; (b) entité qui existe; (c) transition entre deux instants (d) création d'une entité	23
13	Modélisation de l'évolution d'entités à différents niveaux de granularité spatio-temporels (Figures adaptées de [Hornsby et Egenhofer, 2002]) : (a) une géospatial lifeline; (b) un lifeline bead; (c) deux entités vont potentiellement se rencontrer; (d) une suite de déplacements spatio-temporels.	24
14	Modèle <i>ESTDM</i> basé sur les évènements et données raster. A gauche la liste des évènements modélisés sous la forme des différents changements qui s'opèrent sur la carte de base; à droite la façon dont sont représentés ces changements (Figures extraites de [Peuquet et Duan, 1995]).	26
15	(a) Typologie des processus pour une entité selon [Claramunt et Thériault, 1995] : rose, évolution d'une entité ; bleu, relations fonctionnelles entre entités ; jaune, évolution de la structure spatiale d'entités; (b) Taxonomie des processus d'un ensemble d'entités (<i>Set of Geographical Entities SGE</i>) selon [Thériault <i>et al.</i> , 1999]	28
16	Modélisation des changements sous la forme de morphismes d'arbres selon [Jiang et Worboys, 2009] : (a) cas d'une séparation (morphisme surjectif); (b) cas d'une insertion (morphisme injectif)	29
17	Graphe de connectivité modélisant des espaces	30
18	Graphe orienté modélisant l'évolution de parcelle selon une filiation spatiale. . . .	32

19	Modélisation de l'évolution d'entité selon [Stell, 2003] : (a) Evolution d'une famille d'individus, la relation support représente les liens génétiques, P_1 et P_2 sont les parents de P_3 ; Il n'y a pas d'information concernant P_2 en t_3 ; (b) Evolution de parcelles, la relation support représente la superposition (OV), entre autres on a $C OV A$, $C OV B$ et $D OVC$	33
20	Le domaine de temps défini pour le modèle de [Stell, 2003] autorise d'autres représentations du temps que la structure linéaire classique. A gauche on voit une partie du domaine de temps qui évolue en parallèle, à droite les évolutions se rejoignent et se séparent.	36
21	Résumé de la formalisation du modèle spatio-temporel de [Stell, 2003] : chaque lettre majuscule correspond à un concept que nous avons réutilisé tel quel ou adapté dans notre modèle. Cette notation est reprise dans le texte.	37
22	Les relations du $RCC8$	38
23	Représentation de ρ_s à un temps t : à gauche une relation de connexion spatiale entre a et b ; à droite il n'y a pas de relation de connexion spatiale entre ces entités.	39
24	L'espace occupé par a au temps t_i est en relation de connexion spatiale ρ_s avec celui occupé par b au temps t_j , il existe donc une relation spatio-temporelle ρ_{st} entre ces deux entités.	40
25	Représentation de la relation spatio-temporelle entre deux entités : à gauche l'entité a au temps t_i a une relation spatio-temporelle avec l'entité b au temps t_j , à droite a et b n'ont pas de relation spatio-temporelle. Comme pour la relation spatiale, l'arête entre les deux entités en relations peut être étiquetée avec la relation topologique souhaitée.	40
26	Représentation de ρ_f : à gauche, une entité a a une relation de filiation avec l'entité b ; sur la droite, il n'y a pas de relation de filiation entre ces entités.	41
27	Relations de filiations entre deux entités : continuation entre a et b , dérivation entre a et c	41
28	Relations non admises par le modèle : (a) Cas de la division lors d'une relation de continuation; (b) Cas de la fusion lors d'une relation de continuation	41

29	Un exemple de graphe spatio-temporel G_{ST} . Les relations spatiales sont en lignes pleines; les filiations en pointillés avec les étiquettes γ ou δ pour spécifier s'il s'agit respectivement de continuations ou de dérivations, les relations spatio-temporelles sont en lignes doubles.	43
30	(a) cas de transitivité de la relation spatiale; (b) cas de non transitivité de la relation spatiale	43
31	Evolution d'une population de ronces.	44
32	Graphe spatio-temporel G_{ST} associé à l'évolution de la population de ronces.	45
33	Vue schématique des mécanismes de propagation des ronces : (a) drageonnage; (b) marcottage; (c) dispersion des graines; (d) greffe.	46
34	Modélisation des quatre modes de propagations des ronces (a) drageonnage ou marcottage; (b) dispersion des graines; (c) greffe	46
35	Identification des processus de propagation dans l'évolution de la population de ronces : (a) La ronce e est greffée sur la ronce b_1 pour former b_1e ; (b) e_1 est issue soit d'un processus de marcottage ou bien de drageonnage de e ; (c) b_{11} est produite par dispersion des graines de la ronce b_1	47
36	On souhaite que les relations Q et W - et plus généralement R - soient transitives.	49
37	(1) Fonction de voisinage spatial direct du sommet a ; (2) Fonction de voisinage spatial à une certaine distance, ici les voisins de a à la distance 2.	50
38	(1) Fonctions de voisinage spatio-temporel ρ_{st}^- et ρ_{st}^+ ; (2) Fonctions de voisinage spatio-temporel ρ_{st}^{-d} et ρ_{st}^{+d} , $d \geq 1$	51
39	Graphe spatio-temporel G_{ST} associé à l'évolution de la population de ronces. Les entités entourées en noir sont contaminées.	53
40	Union G' de deux graphes spatio-temporels $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_1)$	57
41	Intersection G' de deux graphes $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_1)$	57
42	Exemple de jointure spatiale $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_1)$ tel que $\rho_s[v, x]$, avec $v \in X(t_i) \subseteq V_1$, $x \in X(t_i) \subseteq V_2$	58
43	Exemple de jointure spatio-temporelle entre $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_2)$ tel que $\rho_s[t, y]$, avec $t \in X(t_i) \subseteq V_1$, $y \in X(t_j) \subseteq V_2$, $t_i < t_j$	58

44	Exemple de jointure de filiation entre $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_2,)$ tel que $\rho_f[t, y]$, avec $t \in X(t_i) \subseteq V_1$, $y \in X(t_i) \subseteq V_2$, et ρ_f une relation de dérivation dans Φ	59
45	Exemple de jointure de filiation temporelle, entre $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_2)$ tel que $\rho_f[w, x]$, avec $w \subseteq X(t_i) \in V_1$, $x \in X(t_j) \subseteq V_2$, $t_i < t_j$, et $\rho_f \in \Phi$	59
46	Exemple de jointure de filiation temporelle pour des domaines de temps se chevauchant, i.e. $\exists t_i, t_k \in T_1$ et $\exists t_j, t_l \in T_2$ tels que $t_i < t_j < t_k < t_l$ et $\rho_f[w, x]$ avec $w \in X(t_i) \subseteq V_1$, $x \in X(t_j) \subseteq V_2$ et $\rho_f = \delta$	60
47	Graphe spatio-temporel G_1 de l'évolution des régions. Les relations spatiales et spatio-temporelles sont étiquetées par les relations topologiques du <i>RCC8</i>	61
48	Graphe spatio-temporel G_2 de l'évolution des provinces. Les relation spatiales et spatio-temporelles sont étiquetées par les relations topologiques du <i>RCC8</i>	62
49	Jointure spatiale entre provinces et régions au Chili.	63
50	Exemple de relation entre A et B conforme à la propriété $(A \diamond R B) \wedge \neg(A \boxplus R B) \wedge (B \boxplus R^{-1} A)$. Pour l'illustration on s'appuie sur R définie ici comme relation spatio-temporelle.	65
51	Diagramme des relations d'implications entre les différentes propriétés utilisées pour qualifier les routes existant entre deux entités A et B	66
52	Les différents types de routes	66
53	Représentation des types de routes entre deux entités A et B sous forme de treillis	67
54	Représentation des types de routes entre deux entités A et B sous forme de treillis pour la relation de filiation	68
55	Graphe spatio-temporel G_{ST} associé à l'évolution de la population de ronces.	69
56	Correspondance entre (a) un graphe spatio-temporel et (b) son instanciacion comme schéma de données Σ	79

57	(a) Le graphe spatio-temporel résultat de la jointure entre régions et provinces ; (b) L'instanciation du graphe des provinces comme schéma de données Σ . L'instanciation du graphe résultat de la jointure dans Σ correspond à l'union du schéma de données des provinces (ci-dessus en b) et celui des régions (Figure 56).	81
58	Division de régions	84
59	Séparation de régions.	85
60	Fusion de régions.	85
61	Annexion de régions.	85
62	Trois niveaux de granularité sémantique d'un extrait de la classification végétale de l'ordre des <i>Rosales</i>	93
63	Deux niveaux de granularité temporelle.	93
64	Deux définitions de la granularité spatiale - (1) Deux niveaux de granularité spatiale en fonction de la pente et de l'orientation pour le mont Seorak (Figure issue de [Saux, 2010]) : le critère de pente est pris dans un intervalle de valeurs plus grand en (b) qu'en (a); (2) Trois niveaux de granularité spatiale (et éventuellement sémantique) pour l'étude d'une population : personnes / appartements / immeubles	94
65	Scénario de l'évolution de personnes dans leur lieux de vie, sur le domaine de temps $T = \{T_1, T_2, T_3\}$. Le bâti se compose de deux immeubles à deux étages, chaque étage comporte trois appartements.	97
66	Modélisation par un graphe spatio-temporel d'un scénario présenté Figure 65 sur plusieurs niveaux de granularité spatiale.	98
67	Opérateur d'amalgamation sur l'exemple au temps t_1 : Après passage de cet opérateur sur le graphe spatio-temporel, un niveau de granularité représente un des trois ensembles d'entités décrits sur ce dessin : les personnes, les appartements ou les immeubles suivant le niveau de détail choisi. Les liens (arcs en noirs sur le dessin) entre ces trois types d'entités se retrouvent par l'utilisation de cet opérateur.	99
68	Exemple de bigraphe : A un instant donné l'espace dans lequel sont localisées les entités est identifié par deux <i>sites</i> (0 et 1). On observe par ailleurs deux entités de type <i>immeuble</i> (B_1, B_2), deux entités de type <i>appartement</i> (F_1, F_2) respectivement dans B_1 et B_2 , et trois entités de type <i>personne</i> P_1 et E_1 dans F_1 et E_2 dans F_2 .	102

69	Exemples de représentation d'entités et leurs liens de filiation : (a) Une personne P possède potentiellement 3 ports relatifs à sa mère, son père et ses enfants; (b) Deux personnes P_1 et P_2 et leurs enfants E_1 et E_2 ; (c) Un immeuble B qui ici ne possède pas de lien de filiation	103
70	Exemple de <i>site</i> (0 et 1) et de liens extérieurs (o_{m1}, o_{f1}, o_{f2}).	103
71	Exemple de graphe de liens	104
72	Exemples de graphes de places : (a) possède deux interfaces extérieures et aucune intérieure; (b) possède une interface extérieure et deux interfaces intérieures.	105
73	Représentation des relations spatiales entre entités sous la forme : (a) d'une représentation ensembliste; (b) d'un graphe orienté acyclique - Permet la modélisation de l'ensemble des relations spatiales du <i>RCC8</i> (ici le chevauchement <i>PO</i>), a priori la disjonction ne serait pas matérialisée afin de minimiser le nombre de liens.	105
74	Exemple de bigraphe	106
75	Un exemple de modélisation des différents niveaux de granularité spatiale avec un bigraphe (équivalent au bigraphe de la Figure 74).	107
76	Graphe de places relatif au bigraphe de la Figure 75. On retrouve les sites extérieurs 0 et 1, les sites intérieurs 2, 3, 4 et 5, ainsi que l'ensemble des entités de la représentation équivalente Figure 74.	107
77	Représentation de la règle <i>Entrer</i> : $B.id_0 P_x.id_1 \rightarrow B.(id_0 (P_x.id_1))$	108
78	Représentation de la règle <i>Sortir</i> : $B.(id_0 (P_x.id_1)) \rightarrow B.id_0 P_x.id_1$	108
79	Représentation de la règle <i>Fusion</i> : $E_x H_y \rightarrow (E + H)_{xy}$	109
80	Représentation de la règle <i>division</i> : $E_x \rightarrow F_x G_x$	109
81	Représentation de la règle <i>Capture</i> : $F_x.(id_0 id_2 G_y.(id_1)) \rightarrow F_x.(id_2 G_y.(id_1 H_x.(id_0)))$	110
82	Représentation de la règle <i>Aspiration</i> : $F_x.(id_2 G_y.(id_1 I_z.(id_0))) \rightarrow F_{xz}.(id_0 id_2 G_y.(id_1))$	110
83	Zone d'étude	125
84	Données bathymétriques	126
85	Extraction automatique des paramètres attributaires à l'aide du logiciel ParamDune127	127

86	Représentation des données dans le logiciel ArcGis	127
87	Intégration des données de dunes dans le prototype	128
88	Un exemple de champ de dunes plus complexe (<i>Banc du four</i>) - Image IFREMER	128

Liste des tableaux

I.1	Caractéristiques souhaitées d'un SIG temporel selon [Peuquet et Wentz, 1994]	18
I.2	Caractéristiques nécessaires à la conception d'un modèle spatio-temporel.	19
I.3	Comparatif des différents modèles basés sur la discrétisation de l'espace.	25
I.4	Comparatif des différents modèles basés sur les concepts d'évènement et de processus.	29
I.5	Comparatif des différents modèles basés sur les concepts de théorie des graphes. Les trois derniers modèles ne représentent pas explicitement un mouvement mais sont capables d'extraire un changement spatial entre deux temps.	33

Glossary

Notation	Description	Page
Γ	Ensemble des relations valides pour la relation spatiale	38
Φ	Ensemble des relations valides pour la relation de filiation	40
$\mathcal{P}(X)$	Ensemble des parties de X	48
\mathbb{N}	Ensemble des entiers naturels	38
\mathbb{O}	Ensemble fini des ordinaux	103
σ	Ensemble de toutes les entités sur T	36

Introduction

Contexte

Le développement des sciences de l'information géographique n'a cessé de croître depuis les années soixante-dix, grâce en particulier au développement des Systèmes d'Information Géographique (**SIG**). Ils se sont progressivement imposés comme des méthodes et des outils capables d'organiser, de représenter, d'analyser et de visualiser des données spatialement référencées. Il n'en reste pas moins que l'intégration de la dimension spatio-temporelle est encore l'un des challenges actuels de la recherche en sciences de l'information géographique. Cette dernière repose sur la définition et le développement de modèles (dit modèles spatio-temporels) adaptés au contexte et aux objectifs de l'étude de phénomènes géographiques réels.

Quels que soient ces objectifs, l'analyse de ces phénomènes implique de connaître quels sont les éléments constitutifs qui les caractérisent (le *quoi*), la répartition spatiale de ces éléments (le *où*), et à quel moment ces phénomènes surviennent (le *quand*) [Peuquet et Wentz, 1994]. A ces trois composantes, il faut ajouter la description des processus qui produisent ce (ces) phénomène(s) (le *comment*) [Thériault et Claramunt, 1999]. La problématique induite par ces quatre composantes (*Quoi, Où, Quand, Comment*) fournit les hypothèses de notre recherche et vont nous permettre d'élaborer un modèle capable de représenter et d'analyser un phénomène spatio-temporel.

✧ *Quoi* ?

Pour la représentation de l'espace géographique, [Peuquet, 1984] définit le "*quoi*" comme étant ce que l'on retient du phénomène, c'est-à-dire sélectionner l'information qui sera représentée. Plus généralement, il s'agit d'identifier les différents types d'entités rencontrées, leurs propriétés géométriques (i.e. leur forme) et les attributs permettant d'en calculer la sémantique. Ces entités peuvent être caractérisées par des frontières non clairement définies (entité dite *fiat*, e.g. une

montagne) ou par des frontières connues (entité dite *bona fide*, e.g. un immeuble) [Smith, 1994].

✧ *Où ?*

Où le phénomène se produit renvoie à des informations spatiales de localisation (absolues ou relatives), données par exemple par la distance ou la position relative d'un objet par rapport à un autre (e.g. inclusion, en dehors de). Cette composante est étroitement liée à la caractérisation des relations spatiales entre les entités : relations topologiques ([Egenhofer, 1989, Randell *et al.*, 1992]), métriques ([Pullar et Egenhofer, 1988]) ou d'orientation ([Frank, 1996, Zimmermann et Freksa, 1996, Ligozat, 1998]).

✧ *Quand ?*

La troisième composante, *Quand*, fait naturellement référence à la dimension temporelle ([Allen, 1984, Ladkin, 1987, Freksa, 1992a, Frank, 1994]) dans laquelle s'inscrit le phénomène. L'étude peut être statique s'il s'agit d'un "arrêt sur image" limité à un instant ou une période en particulier, ou dynamique si elle prend en compte des changements se produisant sur un ensemble d'instantanés ou de périodes.

✧ *Comment ?*

Le dernier point, *Comment*, exprime la façon dont un phénomène se déroule, les différents processus mis en jeu et leurs origines. En particulier, l'observation de la dynamique d'un phénomène relève davantage de la compréhension du processus de changement d'état plutôt que de la comparaison d'états successifs [Frank, 2001]. En fait, la modélisation d'un processus de changement d'état et la comparaison d'états successifs sont deux façons complémentaires d'aborder la réponse à la question du *Comment*, dans le cadre de l'étude de la dynamique d'un phénomène.

De la réalité au modèle

Un modèle informatique est une représentation de la réalité contrainte par la nature finie et discrète des systèmes informatiques [Goodchild *et al.*, 1992]. La complexité du monde réel est simplifiée en identifiant les éléments pertinents de ceux qui le sont moins. D'où la nature très subjective d'un modèle qui ne peut être valide pour toute application [Ott et Swiaczny, 2001]. Néanmoins, il reste souhaitable de chercher à élaborer un modèle qui soit le plus **générique** possible, quitte à proposer des adaptations afin de répondre aux exigences propres d'un maximum d'applications. La généralité a cependant ses limites, et certains choix sont garantis du cadre général que l'on souhaite donner au modèle. En particulier, pour ce qui nous concerne, le modèle que nous mettons en

œuvre doit permettre la **représentation, la gestion et la manipulation de données spatio-temporelles**, en faisant abstraction autant que possible des propriétés spécifiques à un cadre d'application précis.

Un grand nombre de recherches ont abordé cette problématique. Elles ont abouti à des modèles basés (1) sur des approches continues ou discrètes ([Armstrong, 1988, Langran et Chrisman, 1988, Worboys, 1992, Egenhofer et Al-Taha, 1992, Hornsby et Egenhofer, 1997, Hornsby et Egenhofer, 2002]), (2) sur des concepts d'évènements et de processus ([Peuquet et Duan, 1995, Claramunt et Thériault, 1995, Thériault *et al.*, 1999, Worboys, 2005]) ou (3) sur des approches basées sur la théorie des graphes ([Jiang *et al.*, 2000, Spéry *et al.*, 2001, Stell, 2003, Jiang et Claramunt, 2004a, Sriti *et al.*, 2005]). Ces derniers travaux ayant suggéré, et montré, l'intérêt offert par la théorie des graphes, nous avons choisi de baser notre approche sur ce type de structure. En effet, elle semble particulièrement adaptée à la représentation, la gestion et la manipulation de relations spatio-temporelles entre entités évoluant dans l'espace. Nous nous plaçons dans la continuité du modèle spatio-temporel générique proposé par [Stell, 2003] qui ouvre la voie à l'élaboration d'une structure de graphe orientée vers la représentation et la **gestion de plusieurs types de relations entre entités** : les relations **spatiales** et celles relatives à l'**identité des entités** à un instant et dans le temps.

Cette modélisation par la théorie des graphes consiste à **représenter** les phénomènes naturels (e.g. évolution de l'organisation de parcelles agricoles, suivi de l'évolution des liens généalogiques entre individus) sous la forme d'un graphe spatio-temporel. La **gestion** des données spatio-temporelles relatives à un phénomène modélisé par cette voie implique (1) de spécifier le modèle de données associé au graphe, (2) de s'assurer de leur cohérence (i.e. consistance) au sein du modèle de données. Une telle condition impose la spécification de contraintes inhérentes à la formalisation du modèle et de celles dépendantes de l'application modélisée. La **manipulation** des données relatives au phénomène est ensuite réalisée au moyen de fonctions génériques, dans le sens où elles peuvent s'appliquer à un grand nombre d'application (e.g. voisinage spatial/spatio-temporel, identification des ancêtres/successeurs d'une entité), et d'autres plus spécifiques (e.g. identification du processus de marcottage dans la propagation d'une population de ronces). Par ailleurs, afin de permettre une analyse approfondie d'un phénomène, ce dernier doit pouvoir être étudié à différents niveaux de détails dans l'espace. L'intégration de fonctions au modèle précédemment défini permet d'aborder cette problématique de la granularité spatiale. Nous étendons cette réflexion à un modèle basé sur les bigraphes.

Finalement, cette recherche orientée vers la représentation, la gestion et la manipulation de données spatio-temporelles, fait appel à plusieurs domaines qui sont présentés Figure 1. Globalement, la théorie des graphes nous apporte les principes et les outils de développement d'une structure de données sur laquelle nous allons baser notre modèle. Le raisonnement spatial et temporel nous fournissent les concepts sur lesquels est bâti le modèle, et les SIG sont le domaine applicatif privilégié pour lequel est construit ce modèle.

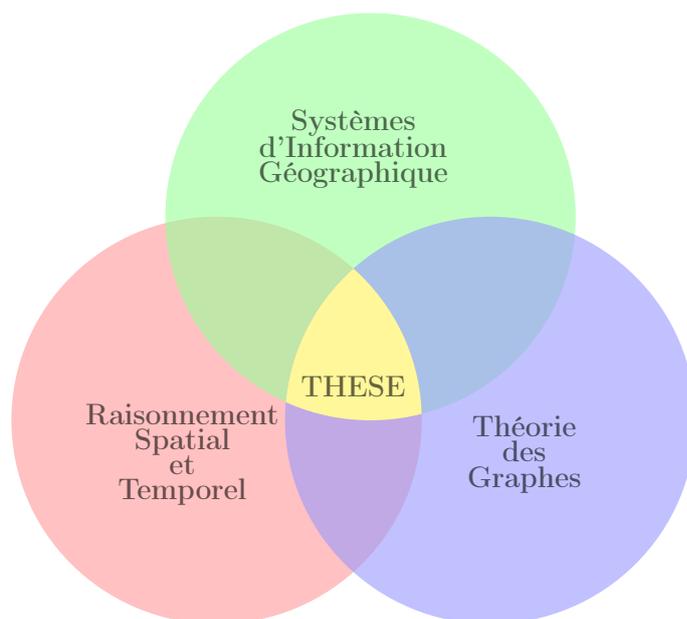


Figure 1 – Contextualisation de la thèse suivant les différents domaines scientifiques abordés

Plan du manuscrit

Le document est organisé selon quatre chapitres dont l'organisation est résumée Figure 2 :

- Le premier chapitre décrit l'état de l'art du domaine des sciences de l'information géographique au travers des concepts d'entité, d'espace et de temps. Un ensemble de modèles spatio-temporels sont détaillés et replacent notre modèle (défini formellement au chapitre deux) dans ce contexte.
- Le deuxième chapitre propose une formalisation de notre modèle sous la forme d'un graphe. Ce chapitre fournit également des outils de description et de manipulation des données spatio-temporelles contenues dans le graphe, au travers d'un ensemble de fonctions. Une illustration de l'utilisation de ces fonctions est proposée par le développement de deux cas d'études relatifs respectivement au phénomène de propagation des ronces et à l'évolution des

régions et des provinces au Chili au début du siècle dernier.

- Le troisième chapitre décrit un modèle de données associé au modèle de graphe défini dans le Chapitre 2. La structuration formelle des données au sein du modèle est détaillée et caractérisée par la spécification des contraintes sur ce modèle. Ces contraintes ont pour objectif de vérifier la cohérence des données au vu de (1) contraintes imposées par notre modèle, (2) contraintes spécifiques à l'application modélisée. Pour le deuxième point, nous proposons quelques exemples destinés à montrer comment spécifier des contraintes relatives à d'autres applications, à partir de ce modèle de données.
- Le quatrième chapitre propose une réflexion prospective qui explore la façon dont on peut gérer la granularité spatiale au sein de notre modèle. Nous proposons une adaptation de notre approche vers un modèle basé sur les bigraphes qui nous semble posséder des caractéristiques facilitant l'intégration de cette composante.

Nous concluons cette thèse en proposant une synthèse des apports et des limites du modèle. Des perspectives de recherche relatives à ce modèle, à son implémentation et à sa validation applicative termineront ce manuscrit.

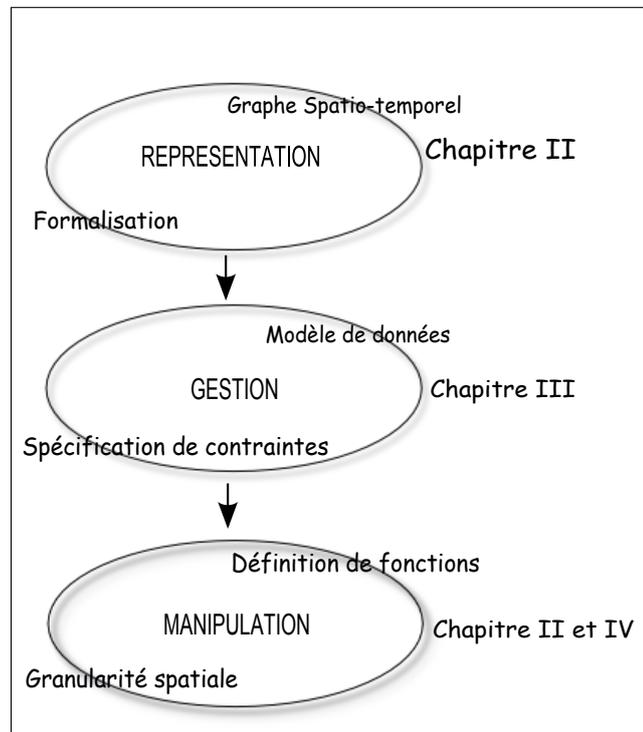


Figure 2 – Méthodologie employée par cette recherche en vue de l'étude d'un phénomène spatio-temporel

I Etat de l'art

Sommaire

I.1	Entité et identité	7
I.2	Modélisation des entités spatio-temporelles	10
I.2.1	Conception de l'espace et du temps	10
I.2.1.1	L'espace	10
I.2.1.2	Le temps	13
I.2.2	Dimension spatio-temporelle	17
I.2.2.1	Les modèles basés sur une approche discrète ou continue	19
I.2.2.2	Les modèles basés sur les concepts d'évènements et de processus	25
I.3	Les graphes comme support de modélisation spatio-temporelle	30

Nous présentons dans ce chapitre le concept d'entité et celui d'identité, ce dernier nous permettant de spécifier les relations non spatiales entre les entités. Notre démarche de modélisation inscrivant ces entités dans l'espace et dans le temps, nous détaillerons ensuite ces deux notions. Finalement, nous terminerons ce chapitre en replaçant notre travail dans le contexte de la recherche en modélisation spatio-temporelle, par la description de modèles proposés dans la littérature.

I.1 Entité et identité

Le déplacement d'une population, la propagation d'un virus, ou la gestion de trafic routier sont des exemples de phénomènes modélisables soit d'un point de vue global quantitatif, par l'analyse des flux du phénomène, (e.g. pour le trafic routier mesurer un débit horaire), soit par des approches discrètes focalisées sur les entités qui les constituent (e.g. pour le trafic routier suivre la trajectoire

d'un véhicule). Nous nous plaçons dans le cadre de la seconde modélisation. Une entité spatiale¹ se caractérise par son empreinte spatiale, son identité et ses éventuels attributs.

Dans les Systèmes d'Information Géographique (SIG), une entité désigne par exemple une parcelle de terrain, une rivière, un bâtiment. Nous définissons l'*empreinte spatiale* de cette entité par l'espace qui lui est directement associé. Mais comment déterminer cet espace ? [Goodchild, 1989, Yuan, 1996] notent que les phénomènes géographiques sont souvent décomposés en objets spatiaux (entités spatiales). Par exemple, les nombreuses études qui se sont penchées sur l'analyse cadastrale utilisent une parcelle de terrain comme entité spatiale de base (e.g. [Spéry *et al.*, 2001], [Rodríguez *et al.*, 2008]). Dans ce cas, l'empreinte spatiale d'une parcelle est représentée par l'aire du polygone associé. De la même façon, un feu de forêt peut être modélisé par un ensemble de polygones (dans cet exemple on peut aussi parler de parcelles), chaque polygone étant différencié en utilisant l'information permettant de déterminer si la zone est brûlée ou non [Langran et Chrisman, 1988]. Par ailleurs, et c'est le cas dans cette thèse, pour une entité mobile, l'empreinte spatiale peut être appréhendée de la même façon en considérant le polygone au temps où les données sont disponibles².

L'empreinte spatiale peut donc être une *surface* mais elle peut aussi être réduite à un *point*, une *ligne* ([Kurata et Egenhofer, 2006, Freksa, 1992b, Schlieder, 1995]), voire étendue à une représentation en 3D ([Kurata, 2009]). Elle dépend théoriquement en partie de la conceptualisation de l'espace. En particulier dans les SIG, il existe principalement deux types de conceptualisation : la vision raster où l'on structure l'espace en un certain nombre de pixels (i.e. l'espace est modélisé par une grille³ et l'empreinte spatiale d'une entité est relative à un ensemble de pixels), et la vision vecteur où l'on représente les données graphiques par des primitives géométriques (i.e. les empreintes spatiales sont relatives à des points, lignes, polygones). [Couclelis, 1992] montre qu'il convient de dépasser le clivage entre ces deux représentations (raster/vecteur) dans la conceptualisation de l'espace géographique. Autrement dit, cette question, raster ou vecteur, se pose souvent en termes techniques (e.g. quel choix d'implémentation ?) alors qu'il faudrait davantage réfléchir en terme de conceptualisation (e.g. cognitivement, qu'est-ce qui est le plus approprié ?). En l'occurrence, du point de vue de cette thèse, l'approche utilisée n'a pas d'importance dans le sens où l'empreinte spatiale d'une entité peut aussi bien être un ensemble de pixels,⁴ qu'une primitive géométrique,

1. Dans cette thèse, nous ne traiterons que d'entités spatiales, par conséquent les termes *entité* et *entité spatiale* sont confondus dans la suite du rapport.

2. Pour une même entité, l'empreinte spatiale peut être différente dans le temps

3. Dans les faits cette grille peut être régulière (pixels) ou irrégulière.

4. On pourrait dire qu'un pixel est aussi une primitive géométrique, la différence majeure réside dans le fait qu'un pixel peut ne représenter qu'une partie d'une entité, alors que la vision vecteur sous entend que la primitive géométrique correspond à l'entité entière.

ce qui laisse toute latitude à la réflexion qu'il faut mener en amont pour identifier, selon le type d'application à modéliser, l'approche à envisager.

Le choix de la modélisation d'une entité peut être très subjectif et dépend de ce que l'on cherche à représenter, des données disponibles et du besoin en termes de relations spatiales à caractériser. Pour une entité de type *bâtiment*, si l'on s'intéresse uniquement à sa localisation géographique, il est possible de la représenter par un point. Par contre si l'on souhaite étudier des relations spatiales impliquant l'empreinte spatiale du bâtiment (e.g. relation de connexion 'touche' pour identifier ses voisins), on préférera la modélisation sous la forme d'un polygone. A contrario, une empreinte spatiale de type ligne est privilégiée dans le cas où l'entité à représenter est un objet à une dimension (e.g. réseau électrique) ou une trajectoire. Dans tous les cas, l'approche dépend de la conceptualisation que s'en fait la personne qui procède à la modélisation.

Caractériser une entité ne se résume pas au choix de son empreinte spatiale. Si la définition d'un phénomène est intimement liée à celle de ses entités, elle est liée aussi à l'ensemble des changements qui sont opérés sur ses entités. Un moyen pour [Hornsby et Egenhofer, 2000, Hornsby et Egenhofer, 1997] de caractériser ces changements est d'utiliser le concept d'*identité* d'une entité. L'identité représente le caractère unique d'un objet, indépendamment de ces attributs ou de ses valeurs. C'est le trait qui distingue un objet de tous les autres [Khoshafian et Copeland, 1986].

L'identité est une notion essentielle pour la conceptualisation et la modélisation d'un phénomène, mais également pour la transcription de l'information dans une base de données. Elle permet de distinguer toute entité d'une autre. Par exemple deux clones ont une identité différente et ce, même si leurs attributs sont identiques [Clifford et Croker, 1988]. Au passage on remarque que le terme d'entité (terme qui représente une abstraction d'un élément du monde réel) est remplacé par *objet* dès lors que l'entité est inscrite dans une base de données [Claramunt et Thériault, 1996]. Dans cette thèse, par abus de langage, et parce que notre approche est davantage orientée modèle, nous n'utiliserons que le terme d'entité quel que soit le contexte.

Cette notion d'identité est très subjective et peut varier selon le type d'entité considéré. Par exemple, dans le cas de personnes, on définit assez naturellement l'identité comme étant les liens entre individus ayant un patrimoine génétique commun, et de ce fait, ces individus appartiennent à une même famille (i.e. filiation). Définir l'identité d'une entité de type parcelle est moins naturel. Une solution peut consister à confondre l'identité de la parcelle avec celle de son propriétaire, une autre avec celle du type de cultures qui s'y trouve. Dans tous les cas, c'est à l'utilisateur de décider ce qui, pour lui, représente l'identité d'une entité, en fonction de son type et du phénomène considéré.

Différencier l'identité de l'ensemble des attributs qui caractérise une entité n'est pas toujours facile, et ce, même si d'un point de vue conceptuel ils sont indépendants. Les attributs d'une entité sont les éléments qui la caractérisent. Les attributs d'une personne peuvent être son nom, âge, couleur des yeux etc. Dans l'exemple du feu de forêt de [Langran et Chrisman, 1988], l'information qui précise si la parcelle est brûlée ou non constitue un attribut de l'entité parcelle. Dans les faits, au niveau des bases de données, l'identité est souvent confondue avec l'identifiant unique de l'entité. C'est donc finalement l'un de ses attributs. En ce qui nous concerne, nous nous efforcerons de nous appuyer sur la définition conceptuelle de l'identité, qui différencie l'identité d'une entité de l'ensemble de ses attributs, lorsque nous aurons à la spécifier pour une entité.

Une entité spatiale se caractérise par : son empreinte dans l'espace, son identité et ses attributs. Elle peut être statique (e.g. une parcelle de terrain), ou mobile (e.g. une personne en déplacement).

I.2 Modélisation des entités spatio-temporelles

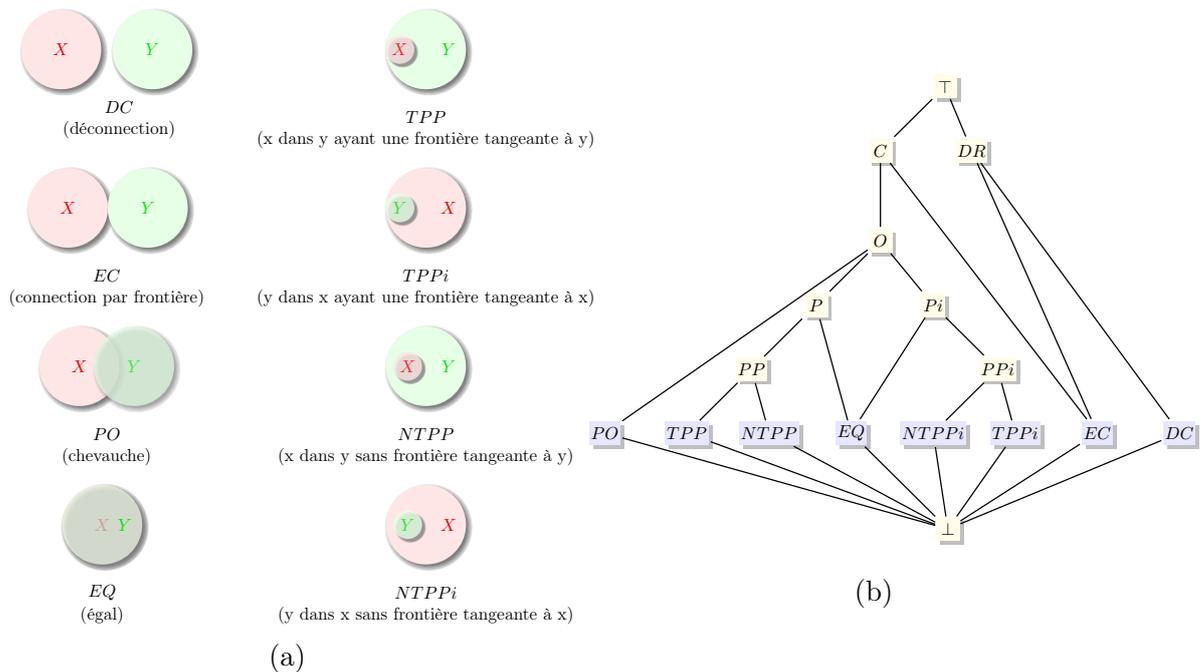
I.2.1 Conception de l'espace et du temps

I.2.1.1 L'espace

Une entité se caractérise fondamentalement par son empreinte dans l'espace, il est donc naturel de penser que les relations entre entités sont liées aux relations spatiales entre leurs empreintes spatiales. Ces relations spatiales se déclinent principalement en trois catégories : topologique, métrique et d'orientation.

Après le travail de [Egenhofer, 1989], les relations spatiales topologiques ont été décrites dans le *Region connection calculus (RCC8)* de [Randell et al., 1992] sous la forme de 8 relations topologiques (Figure 3a). La représentation sous forme de treillis de ces relations permet, en les ordonnant, de caractériser les liens de dépendance qui existent entre elles (Figure 3b).

En plus des relations spatiales entre des entités de types polygones, d'autres types d'empreintes spatiales ont également été décrites. [Schlieder, 1995, Freksa, 1992b, Kurata et Egenhofer, 2006]

Figure 3 – (a) Relations topologiques du *RCC8*; (b) Treillis des relations topologiques

expriment par exemple les relations topologiques entre deux segments de lignes. En particulier, des 8 cas identifiés dans le *RCC8*, [Kurata et Egenhofer, 2006] passent à 12 classes de relations topologiques dans le cadre des entités de ce type (segment de ligne).

[Kurata, 2009] considère les relations topologiques d'entités modélisées par des points, des surfaces ou des lignes et il traite également de la modélisation 3D (ce dernier point a aussi été traité par [Billen et Clementini, 2006]). Il décrit des relations topologiques entre des entités modélisées par des empreintes spatiales de différents types en utilisant la caractérisation de [Egenhofer et Al-Taha, 1992] présentée section I.2.2.1 et rappelée brièvement ici en Figure 4.

On peut remarquer que toutes les relations du *RCC8* sont des connexions, à l'exception de la relation *DC* de disjonction. Il est possible de préciser la disjonction par d'autres spécifications des relations spatiales. Ainsi, [Pullar et Egenhofer, 1988] qualifie une relation avec des notions de métriques (à la distance d), par des concepts plus vagues (à côté, très près, loin, proche), et des indications de direction (points cardinaux). En particulier [Worboys, 2001] a effectué des études sur la notion de proximité, où il demande à plusieurs panels de personnes ce que signifie pour eux la notion de 'proche'. Les relations d'orientation ont été approfondies par [Frank, 1996],

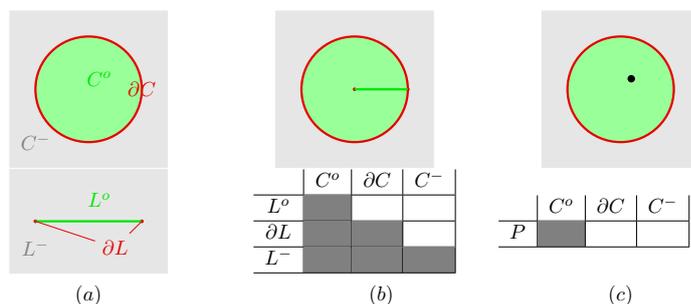


Figure 4 – Exemples de relations topologiques entre deux entités dont le type d’empreinte spatiale est différent (surface et ligne/point) - Figure adaptée de [Kurata, 2009] : (a) spécification de l’extérieur E^- , intérieur E^o , frontière ∂E pour chaque entité ; (b) cas d’une surface C et d’une ligne L , une case de la matrice est grisée s’il existe une relation topologique de connexion ; (c) cas de la surface et du point P qui ne possède ni intérieur, extérieur ou frontière, on identifie uniquement la partie de la surface avec laquelle il est en relation topologique

[Zimmermann et Freksa, 1996] et [Ligozat, 1998].

Quelle que soit la modélisation des relations spatiales choisie, le raisonnement qualitatif qui l’accompagne est primordial dans l’étude des propriétés spatiales relatives aux empreintes spatiales des entités qui constituent un phénomène. Il intervient en particulier à partir du moment où l’on manque de précision sur les données quantitatives disponibles. C’est à dire que l’on n’est pas capable de situer géographiquement où se trouve une entité, mais on peut la situer par rapport à une autre. Ce raisonnement, qui peut être basé sur le *RCC8* ([Cohn et al., 1997, Cohn et Varzi, 1999]) ou sur des relations de cardinalité ([Pullar et Egenhofer, 1988, Ligozat, 1998]), permet de caractériser et d’étudier formellement une situation spatiale donnée. [Cohn et Renz, 2007] ajoutent que le challenge du raisonnement spatial qualitatif (RSQ) est de fournir des systèmes de calcul permettant à une machine de représenter et de raisonner sur des entités spatiales sans faire appel aux techniques traditionnelles (i. e. techniques quantitatives). Dans ce type de raisonnement, la composition des relations est essentielle, d’où l’intérêt de modéliser ces compositions et d’établir les tables correspondantes.

Les relations spatiales entre entités sont de différentes natures (topologiques, cardinales, métriques). Le Raisonnement Qualitatif Spatial est l’outil théorique essentiel basé sur ces relations, permettant l’analyse des propriétés spatiales relatives aux entités spatiales qui constituent un phénomène.

I.2.1.2 Le temps

L'analyse de l'évolution d'entités implique d'étudier le temps comme une composante du phénomène considéré [Bordin, 2010]. Le temps, comme l'espace, peut être structuré et ordonné pour être analysé. [Langran et Chrisman, 1988] montrent que la dimension temporelle est topologiquement similaire à la dimension spatiale. Ils expliquent que la connectivité de l'espace dans une analyse temporelle est aussi importante que la connectivité temporelle dans l'analyse spatiale. Dans le domaine de l'information géographique en particulier, le temps est indissociable de la dimension spatiale et, selon [Parkes et Thrift, 1980], c'est même ce qui distingue la géographie de la géométrie.

Le temps possède plusieurs interprétations possibles ([Snodgrass et Ahn, 1986, Barrera *et al.*, 1991, Jensen *et al.*, 1994]) : valide, transactionnel ou propre à l'utilisateur. Le temps dit *valide* est celui du monde réel, qui est différent du temps *transactionnel*. Ce dernier est défini par un utilisateur, c'est par exemple la date à laquelle un élément est inséré dans une base de données (peut être différent de la date à laquelle il s'est effectivement produit). Le temps *propre à l'utilisateur* réfère à des dates ou des événements qui dépendent de l'utilisateur, il s'agit par exemple d'attributs de date comme "anniversaire" ou "début de vacances". Lors de la définition d'un modèle temporel, on peut gérer la ou les dimensions du temps qui nous intéressent, par exemple, [Worboys, 1994] propose un modèle bi-temporel qui peut prendre en compte des informations reposant sur le temps valide et sur le temps transactionnel dans les SIG.

Quand la dimension est choisie, on peut représenter le temps selon de multiples structures ([Worboys, 1990]) suivant le phénomène considéré : cyclique, parallèle ou linéaire. La modélisation des marées repose sur une structure *cyclique*, alors qu'un ordinateur capable d'effectuer différents calculs simultanément nécessite une représentation du temps *parallèle*. L'évolution d'un groupe d'individus se modélise plutôt de façon *linéaire*, mais elle peut évoluer en temps parallèle ("*branching time*") si le groupe se scinde en plusieurs sous groupes, ce qui renvoie à plusieurs scénarios de simulation, c'est à dire plusieurs évolutions différentes pour un même groupe d'individus (Figure 5).

Classiquement la représentation linéaire est souvent privilégiée. Le temps se modélise sous la forme d'une droite orientée, sans début ni fin pour simuler la dimension infinie du passé et du futur. On y place les états temporels que l'on souhaite matérialiser sous la forme d'instantanés (i.e. état qui n'a pas de durée), concrètement des points sur la droite, ou d'intervalles (e.g. période).

[Galton, 2000] propose une définition de ces états temporels selon leur dénomination. Il distingue

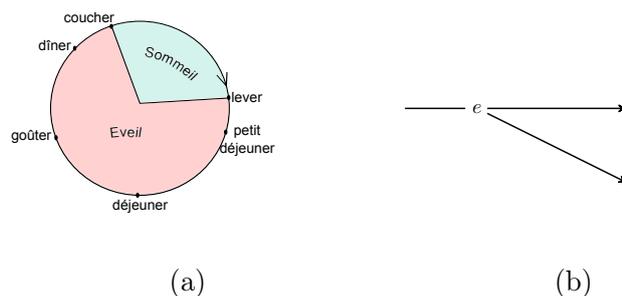


Figure 5 – Structures de temps envisageables selon le type de phénomène à représenter (a) temps cyclique ; (b) le temps se divise en deux ("branching time") à l'instant e , par conséquent des événements peuvent se produire en parallèle ensuite

les *événements* (terme employé par [Allen, 1984]) qui sont des occurrences discrètes d'un phénomène avec un début et une fin (e. g. une randonnée). Ces événements sont constitués de *processus* (e.g. marcher, pleuvoir) qu'il différencie des *routines* (e.g. faire un café, éteindre l'ordinateur). La frontière entre événement et routine peut paraître ténue, dans les faits, une routine devient un événement dès lors qu'elle devient une action exceptionnelle (e.g. préparer le déjeuner est une routine, mais préparer le déjeuner de Noël est un événement).

L'analyse temporelle de phénomènes nécessite de formaliser les relations que l'on peut observer entre états temporels. En particulier, [Allen, 1984] a apporté une contribution pour la représentation des relations temporelles sous la forme d'intervalles. Pour lui, la représentation des événements par des instants sur la droite du temps est source d'inconsistance. Il considère que, par définition, un instant a une durée nulle, alors que d'un point de vue cognitif, un événement doit avoir une certaine durée pour être perceptible dans le temps. C'est pourquoi, il définit une algèbre de 13 relations de bases entre deux intervalles temporels⁵ (Figure 6). Dans [Allen et Hayes, 1985] ils distingueront explicitement les événements qui ont une certaine durée (e.g. des vacances) de ceux qui peuvent être abstraits à un instant (e.g. un click de souris d'ordinateur). Et dans [Allen et Hayes, 1990], les auteurs montrent que finalement une algèbre sur les instants peut tout à fait se dériver de celles des intervalles.

[Ladkin, 1987] complète l'algèbre d'Allen en qualifiant les relations entre des intervalles non convexes. Si on compare un intervalle non convexe i avec un autre intervalle non convexe j , on peut

5. On peut rapprocher cette définition temporelle de celle que proposera plus tard [Randell *et al.*, 1992] pour l'espace avec le *RCC8*

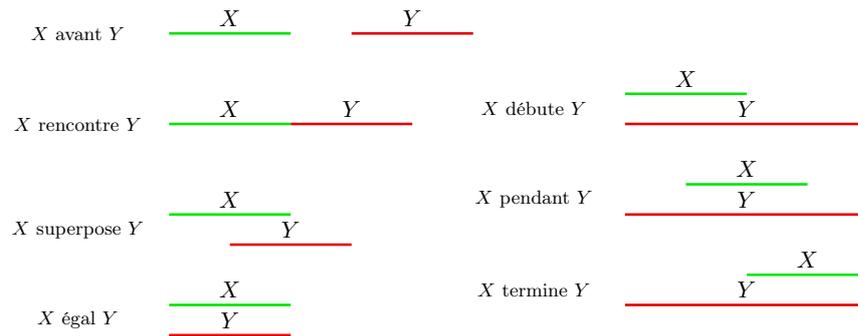


Figure 6 – Algèbre temporel de Allen : à l'exception des intervalles égaux, toutes les relations de X vers Y peuvent être établies du point de vue de Y (i.e. X avant Y , Y après X), d'où les 13 relations

qualifier la relation en disant par exemple que i rencontre j *souvent, toujours, de temps en temps* (5 qualificatifs possibles sont décrits).

Par la suite, [Freksa, 1992a] généralise l'approche proposée par Allen avec les semi-intervalles, qui correspondent à la définition de "nest" (i.e. permettent d'identifier le début et la fin d'évènements très courts, comme la fermeture d'une porte) dans [Allen et Hayes, 1985]. En plus des relations de [Allen, 1984], il prend en compte les relations de voisinages entre les relations temporelles. Les évènements ne sont plus analysés indépendamment mais en fonction des autres évènements.

Le choix de la modélisation du temps (instants/intervalles) subit des contraintes analogues au choix de l'empreinte spatiale. Ce choix dépend de ce que l'on veut représenter, des données disponibles et des besoins en termes de relations temporelles à représenter. Pour une entité de type bâtiment, le temps valide relatif au bâtiment est compris entre sa construction et sa démolition. Si les requêtes ne portent que sur la présence ou non de cette entité à des dates précises, on optera par exemple pour une représentation du temps avec des instants. Si au contraire la comparaison implique de comparer l'existence de deux entités sur différentes périodes (sur quel laps de temps deux bâtiments ont-ils coexisté ?), la modélisation par des intervalles est favorisée.

Bien que le concept de temps doive par définition avoir une représentation continue, sa transcription informatique impose dans les deux cas (i.e. modélisation sous forme d'instants ou d'intervalles), et de manière quasi systématique, sa discrétisation. Concrètement, on est contraint de choisir la granularité temporelle à laquelle on représente le phénomène (e.g. par jour, par année). Il est en effet rare de disposer de toutes les données (sans interruption) sur une période. Néanmoins, si la densité des informations temporelles est suffisante, on considère que l'on dispose d'une certaine continuité. Il faut cependant être conscient du côté subjectif et imparfait de cette discrétisation (que se passe-t-il pendant la milliseconde qui n'est pas prise en compte dans les données ?) et savoir

évaluer son impact lors de l'analyse d'un phénomène.

De plus, on ne connaît pas toujours les dates précises auxquelles se situent les différents évènements, aussi [Frank, 1994] propose une approche qualitative dans laquelle les évènements sont rangés selon l'ordre dans lequel ils apparaissent, sans préciser de données quantitatives. Ainsi, il s'abstrait de la représentation du temps linéaire classique en représentant les évènements sous la forme d'un diagramme de Hasse (Figure 7), et propose un langage formel de Raisonnement Qualitatif Temporel (RQT). Ce dernier, de la même façon que le Raisonnement Qualitatif Spatial pour la dimension spatiale, formalise et permet l'analyse de situations temporelles à partir de l'ordonnement des évènements (e.g. avant, après).

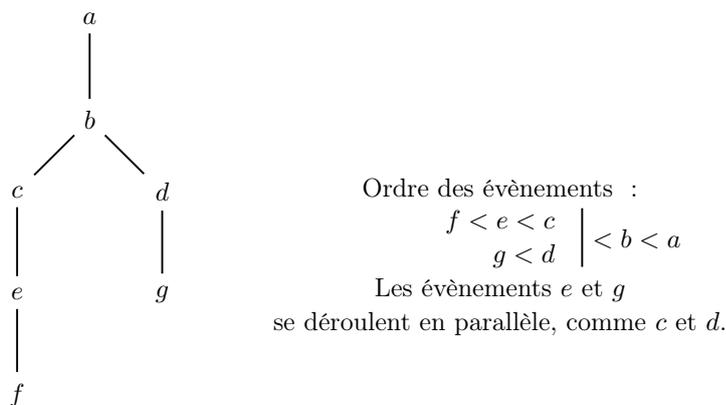


Figure 7 – Diagramme de Hasse pour l'ordonnement qualitatif d'évènements ([Frank, 1994])

Il existe principalement deux modes de représentation du temps, sous la forme d'instants ou d'intervalles, et l'on range communément les différents évènements selon leur date/période d'apparition sur une droite temporelle. Le raisonnement qualitatif temporel permet de s'abstraire de ces données quantitatives, et comme le Raisonnement Qualitatif Spatial pour la dimension spatiale, il formalise et permet l'analyse de situations temporelles.

La modélisation du temps a vocation à être intégrée dans un modèle temporel de données. [Worboys et Hornsby, 2004] définissent un modèle conceptuel comme une représentation structurée d'une partie du monde qui peut être capturée par un système d'information (géographique). S'il en est l'élément clé, le temps n'est pas le seul constituant d'un tel modèle. [Langran, 1992] identifie les trois éléments essentiels d'un modèle de donnée temporel pour les SIG : des *états* (la distribution des entités spatiales), des *évènements* ou *processus* (ce qui provoque les éventuels changements dans la distribution spatiale des entités) et l'*evidence* (comment ces changements sont découverts et mesurés). Un peu plus tard, [Allen et Ferguson, 1994] différencient les *actions* des *évènements*, l'action étant vue comme plus ponctuelle.

Ceci nous montre que la dimension temporelle est indissociable de la dimension spatiale pour l'analyse de phénomène et qu'il faut donc proposer des modèles qui prennent en considération ces deux dimensions.

I.2.2 Dimension spatio-temporelle

Pouvoir élaborer des prédictions à partir d'un ensemble de données représentant un phénomène, ou réduire notre ignorance sur ce phénomène (i.e. accroître l'information dont on dispose) sont selon [Couclelis et Liu, 2000] les principales raisons pour lesquelles nous construisons des modèles. [Peuquet et Wentz, 1994, Peuquet, 2002] élaborent un triangle (TRIAD) qui décrit trois angles de vues simultanées d'une modélisation spatio-temporelle, à savoir la vue fondée sur la position (where), celle fondée sur le temps (when) et la vue fondée sur l'objet (what). Ce sont là les concepts fondamentaux qu'un modèle de données spatio-temporel se doit de pouvoir représenter au mieux.

L'objectif de ces approches spatio-temporelles est de pouvoir représenter et analyser l'ensemble des changements qui sont opérés sur les entités spatiales inhérentes à un phénomène. Dès 1994, [Peuquet et Wentz, 1994] spécifient les besoins d'un SIG temporel (Tableau I.1) et confirment

que l'analyse d'un phénomène passe à la fois par la gestion du temps et de l'espace. Bien que [Yuan, 1996] précise cette liste en catégorisant les différents changements qui peuvent s'opérer au sein de l'information géographique (i.e. statique, transitionnel, mutation, mouvement), nous nous appuyons en partie sur les axes plus généraux dégagés par Peuquet pour comparer les différentes approches spatio-temporelles décrites dans la littérature. Nos critères sont détaillés dans le Tableau I.2. Tous les tableaux comparatifs (i.e. Tableaux I.3, I.4 et I.5) sont renseignés avec quatre qualificatifs - -, -, + et ++ selon que le modèle considéré satisfait plus (+) ou moins (-) les critères. Notons que le qualificatif ++ (respectivement - -) ne signifie pas qu'un modèle est parfait (respectivement très mauvais) dans l'absolu pour ce critère, mais seulement qu'il remplit les exigences attendues dans le cadre de cette évaluation. De plus, si l'on considère qu'un critère est hors de propos pour un modèle (e.g. voir la remarque ci-dessous sur le critère 'mouvement' pour les modèles basés sur l'espace), on indique simplement **NR** pour Non Renseigné.

Le critère relatif au mouvement concerne essentiellement les approches basées sur l'entité (i.e. approches discrètes). On considère que dans les approches basées sur une modélisation continue de l'espace (i.e. *location based*, aussi appelée "approche raster") l'espace est divisé en un pavage régulier ou non (une grille), et cette notion d'entité n'est pas prise en compte. Les valeurs des attributs qui qualifient les différents pavés de cette grille peuvent varier (changement d'état), mais la position de ces pavés dans l'espace ne peut changer au cours du temps : l'espace qu'ils définissent ne peut pas se déplacer. Par conséquent, lorsque par exemple dans le tableau récapitulatif nous choisissons d'y inclure le modèle snapshot, il s'agit du modèle snapshot qui intègre la vision entité et non pas continue.

Pour ce qui concerne la granularité, nous développerons ce concept dans le chapitre IV, nous ne nous étendons donc pas ici sur sa définition et ses caractéristiques. Par ailleurs, certains critères comme la gestion des données sont assez subjectifs. En particulier lorsqu'il s'agit d'extrapoler l'implémentation d'un modèle, qui, dans les faits, n'est que théorique.

-
- Représenter efficacement les changements à travers le temps.
 - Représenter un même espace géographique selon différentes séquences de temps.
 - Faire varier la granularité temporelle (année, jour etc.).
 - Manipuler et analyser des données spatio-temporelles grâce aux relations temporelles (et spatiales).
 - Proposer une interface utilisateur ergonomique et flexible.
-

Tableau I.1 – Caractéristiques souhaitées d'un SIG temporel selon [Peuquet et Wentz, 1994]

Représenter les changements liés à/au :	○ Mouvement	Le modèle permet-il d'identifier/représenter un mouvement local ? e.g Déplacement d'entités entre deux temps.
	○ Changement d'état	Le modèle permet-il d'identifier/représenter un mouvement global ? e.g Résultat qualitatif ou quantitatif décrivant le déplacement d'une entité sur un ensemble de temps. Peut-on identifier qu'un attribut d'une entité (de l'espace) a été modifié entre deux temps ? Peut-on identifier qu'un attribut d'une entité (de l'espace) a été modifié sur plusieurs temps ?
	○ l'Identité	Peut-on savoir si l'identité d'une entité a été modifiée ? Peut-on identifier/représenter/suivre les filiations associées ?
Gestion Granularité :	○ Temporelle	Peut on extraire/analyser les données d'un phénomène à différents niveaux de détails temporels ? Est-ce prévu par le modèle ? Existe-t-il des opérateurs spécifiques ? e.g Passer d'une représentation par heure à une représentation par jour pour un même phénomène
	○ Spatiale	Peut on extraire/analyser les données d'un phénomène à différents niveaux de détails spatiaux ? e.g Trois niveaux de granularité spatiale : Personnes, appartements où vivent ces personnes, et quartiers où sont situés ces appartements
Manipuler et analyser des données ST à partir de relations temporelles et spatiales		Peut-on gérer des données dans le temps et l'espace ? Existe-t-il des opérateurs spécifiques (locaux, globaux) d'analyse de données spatio-temporelles ? Peut-on identifier/analyser des processus et leur causes ?
Gestion des données :	○ Initialisation	L'introduction des données dans le modèle est-elle fastidieuse/optimisée ?
	○ Mise à jour	La mise à jour des données dans le modèle est-elle fastidieuse/optimisée ?
	○ Stockage	Le modèle nécessite -t-il un espace de stockage important ou réduit ?

Tableau I.2 – Caractéristiques nécessaires à la conception d'un modèle spatio-temporel.

I.2.2.1 Les modèles basés sur une approche discrète ou continue

Le modèle *snapshot* ([Armstrong, 1988]) décrit des états spatiaux ordonnés à différents temps sans lien explicite entre eux. A chaque temps correspond une couche (état) indépendante, c'est-à-dire qu'un changement n'est pas explicitement représenté, il faut pour cela comparer deux états différents. Suivant les versions du modèle snapshot, chaque état peut représenter un ensemble temporellement homogène d'unités spatiales (vision continue de l'espace) ou hétérogènes d'entités géographiques (vision discrète de l'espace). Dans les deux cas, l'intervalle de temps entre chaque état peut varier sans incidence sur la cohérence du modèle (Figure 8). Ce modèle permet de fournir des informations quantitatives au niveau global (e. g statistiques concernant l'occupation du sol). En plus de l'absence

de lien temporel, un inconvénient à ce modèle est la redondance d'informations qu'il peut susciter. En effet, si l'on capture des données à intervalles réguliers, il y aura très probablement des états identiques. Même si l'on contraint la création d'un nouvel état seulement à chaque changement qui s'opère sur l'espace, cet état contient à la fois les espaces sur lesquels s'est produit le changement mais aussi tous les autres : la redondance d'information reste donc importante. Par ailleurs, il est difficile d'évaluer les conditions du passé car il est rare qu'un phénomène puisse être caractérisé en prenant en compte une seule couche d'information indépendamment.

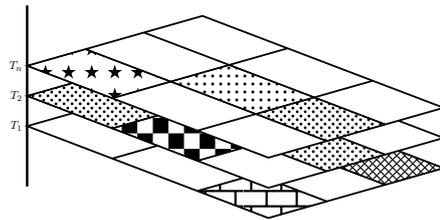


Figure 8 – Le modèle snapshot selon [Armstrong, 1988]. Dans cet exemple, une entité est une parcelle de terrain possédant une certaine propriété (e.g. type de la culture qui s'y trouve)

Le modèle *Space-Time Composites* (STC) ([Langran et Chrisman, 1988]) peut être vu comme une "amalgamation" des états du modèle snapshot. C'est un modèle dans lequel les changements sont "codés" dans les attributs des entités. Un phénomène y est décrit comme un ensemble d'espaces homogènes (objets) sur lesquels s'appliquent les valeurs d'un attribut (un seul attribut par espace). Ces objets sont le résultat de l'amalgamation entre différents fragments d'entités, par croisement de l'ensemble des couches. La suite des changements sur ces objets est incluse dans leurs attributs et capturée à des temps discrets (là aussi, la durée de l'intervalle entre deux temps a peu d'importance, et peut varier) (Figure 9). Outre une mise à jour fastidieuse et la nécessité de disposer d'un espace mémoire de stockage conséquent, si ce modèle peut capturer dans le temps un grand nombre d'attributs, il a pour principal inconvénient de ne pas pouvoir repérer explicitement des changements dans l'espace (e.g. déplacement, superposition).

Le modèle *Spatio-temporal Object* (ST-object) ([Worboys, 1992]) représente le monde réel par un ensemble discret d'objets (entités) qui intègre la dimension temporelle de façon orthogonale à l'espace en deux dimensions. Chaque entité est composée d'*atomes spatio-temporels* (Figure 10),

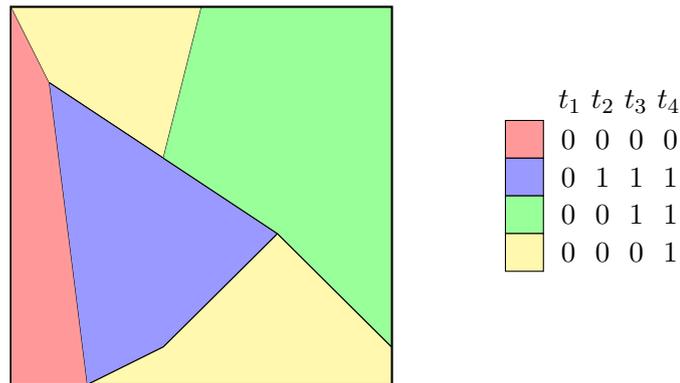


Figure 9 – Exemple de modèle Space-Time Composites modélisant un feu de forêt : la matrice décrit l'évolution du feu dans le temps, 0 (resp. 1) indique une zone non brûlée (resp. brûlée) au temps t considéré - Figure adaptée de [Langran et Chrisman, 1988]

c'est-à-dire les empreintes spatiales de l'entité aux temps auxquels elle existe. Elle est associée à un identifiant, on peut donc vérifier sa persistance dans le temps. Ce modèle est capable de repérer les changements dans les attributs d'une entité, à la fois en spatial et en temporel (ou séparément), et ce, en projetant ses atomes sur la (les) dimension(s) souhaitée(s). Cependant, ce modèle ne représente pas explicitement les changements (e.g. déplacement) et l'accès et l'interrogation de données restent complexes [Yuan, 1999].

De plus, ce modèle, comme les deux précédents ne représente que les changements soudains sur une structure de temps discrète et linéaire, et on ne peut identifier de processus. Aucun de ces modèles ne permet de mettre en évidence les concepts de mouvement ou de changements liés à l'identité d'une entité. Ces modèles ne répondant pas pleinement aux caractéristiques souhaitées (Tableau I.3), d'autres modèles ont vu le jour, apportant des modélisations spécifiques suivant leur définition propre de la notion de changement dans l'évolution d'une entité.

Le modèle de [Egenhofer et Al-Taha, 1992] considère qu'un changement quantitatif apparaît si la déformation (i.e. translation, rotation, réduction, expansion) d'une entité affecte les relations topologiques qu'elle entretient avec d'autres entités. Par exemple lorsqu'un fleuve sort de son lit, sa relation avec les terrains situés juste à côté peut être affectée. Cette modélisation formalise la relation topologique qui existe entre deux entités sous la forme d'une matrice (*g-intersection*). Pour la construire, il définit une entité A comme possédant un *intérieur* (A^o), une *frontière* (∂A) et un *extérieur* (A^-). L'entité B ayant la même spécification, l'intersection des deux donne la matrice suivante, dont les deux seules valeurs sont une intersection vide (\emptyset) ou non vide ($\neg\emptyset$) :

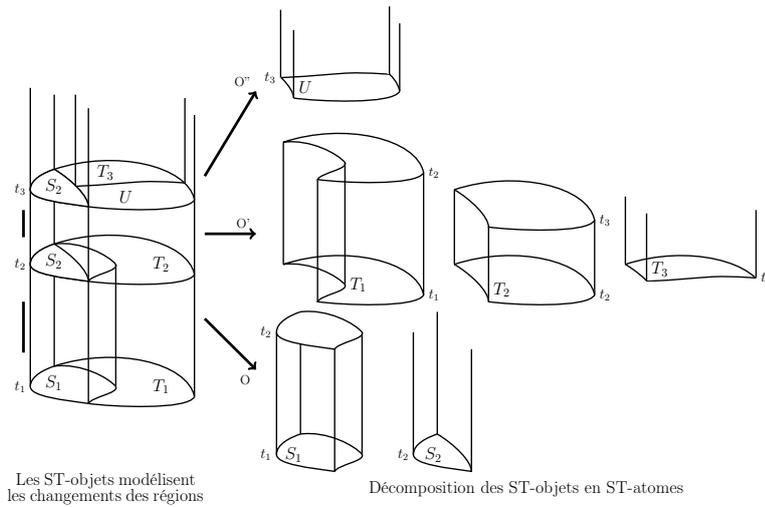


Figure 10 – Le modèle Spatio-temporal Object selon ([Worboys, 1992]). Exemple d'interprétation : l'entité S_1 peut être décomposée en deux atomes qui représentent spatialement l'entité en t_1 et t_2 .

$$\begin{pmatrix} A^o \cap B^o & A^o \cap \partial B & A^o \cap B^- \\ \partial A \cap B^o & \partial A \cap \partial B & \partial A \cap B^- \\ A^- \cap B^o & A^- \cap \partial B & A^- \cap B^- \end{pmatrix}$$

Si entre deux instants la matrice est modifiée, c'est qu'un changement a eu lieu. [Egenhofer et Al-Taha, 1992] ont dénombré 512 matrices possibles, dont celles relatives aux relations topologiques dans \mathbb{R}^2 , c'est-à-dire les 8 correspondantes au *RCC8*. Trois exemples sont proposés Figure 11. Si l'on considère trois temps consécutifs, on voit clairement à chaque étape, une modification de la matrice associée. [Egenhofer et Al-Taha, 1992] a quantifié la *distance* qui sépare une relation topologique d'une autre, afin d'être en mesure de quantifier le changement qui s'est opéré.

Mais ce modèle ne peut pas détecter les mouvements d'entités si elles n'ont pas de relations spatiales entre elles (e.g. deux entités disjointes qui se rapprochent sans se toucher). De plus, la représentation globale de l'évolution des entités est difficile (i.e. vision matrices des relations topologiques, suivre une entité en particulier est compliqué).

Une autre composante entre en jeu pour modéliser l'évolution des entités dans le temps et l'espace : l'identité d'une entité. [Hornsby et Egenhofer, 1997, Hornsby et Egenhofer, 2000] considèrent en

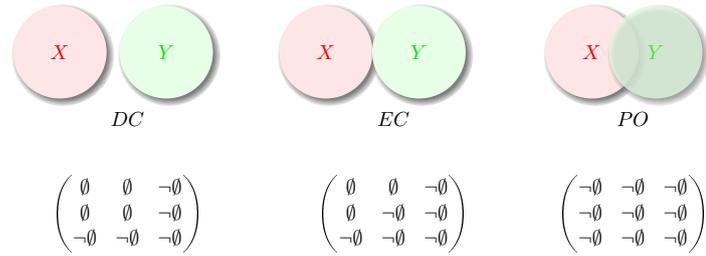


Figure 11 – Quelques relations du *RCC8* vu par [Egenhofer et Al-Taha, 1992]. Si l’on considère ces trois situations à trois temps consécutifs, l’évolution des relations spatiales entre *X* et *Y* se lit dans les matrices associées.

effet un changement s’il y a des modifications de l’identité des entités. Ils mettent au point le langage CDL (Change Description Language) graphique, pour modéliser explicitement les changements (Figure 12). Ce langage se base sur le concept d’existence (présence physique de l’entité) ou de non existence d’une entité qui est modélisée grâce à l’intégration de l’identité dans cette approche. Le modèle sous-jacent est basé sur une modélisation qualitative du temps. Ils élaborent des opérateurs qui agissent sur une entité (création, destruction, continuation), ou sur des transitions (séparation, fusion), qui pour partie ont été définis dans [Medak, 1999] mais par langage formel et non iconographique.

Outre le fait qu’on ne puisse représenter avec le langage *CDL* que certaines informations telles que la durée du changement, et que ce langage ne permette pas de raisonner facilement sur les changements, ce modèle et celui de [Medak, 1999] ont l’inconvénient majeur de ne pouvoir représenter que les changements liés à l’identité [Haddad, 2009].

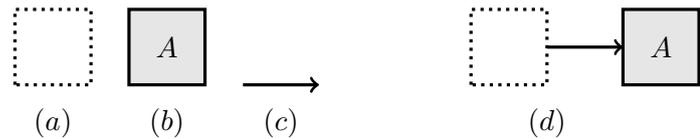


Figure 12 – Quelques primitives de base du langage CDL [Hornsby et Egenhofer, 2000] : (a) entité qui n’existe pas ; (b) entité qui existe ; (c) transition entre deux instants (d) création d’une entité

Il est possible de suivre l’évolution spatiale des entités avec des modèles mettant en œuvre des trajectoires d’entités dans l’espace et le temps. En effet, [Hägerstrand, 1970], père de la *Time Geography* qui justifie l’égale prise en compte de l’espace et du temps dans l’étude des dynamiques spatiales, propose un modèle où les trajectoires sont décrites sous la forme de lignes associées à une activité afin de représenter le déplacement d’une entité. Ces trajectoires permettent une bonne visualisation du mouvement des entités dans l’espace. Ce type de représentation est repris et étendu

par le modèle de [Hornsby et Egenhofer, 2002] au travers des *geolifelines* (les géolifelines sont aussi utilisées par [Thériault *et al.*, 2002] pour la modélisation de trajectoires de vies). Cette approche définit des périodes de temps pendant lesquelles une entité occupe un espace. L'évolution de cette entité pendant cette période est modélisée par des *geospatial lifelines* (l'association déplacement-activité de [Hägerstrand, 1970]) de différentes formes selon le type de mouvement considéré (discret ou continu) (Figure 13a). Un mouvement se décompose en une série de cônes géométriques ou *lifeline bead* (Figures 13b et 13d) qui représentent le déplacement maximal pouvant s'opérer sur une entité à une vitesse maximale donnée. Si l'on observe plusieurs entités, il est possible de déterminer si elles peuvent potentiellement se rencontrer ou non, c'est-à-dire si leurs cônes respectifs ont des points communs (Figure 13c). On peut aussi détecter des différences de vitesses entre les étapes d'un mouvement (différentes formes de cônes), ou s'il ne se passe rien entre deux étapes (le cône est alors réduit à une ligne reliant potentiellement deux cônes (Figure 13d)).

Par ailleurs, ce modèle permet d'observer les changements qui peuvent s'opérer sur une entité à différents niveaux de granularité (i.e. niveaux de détails) spatiale et temporelle, mais les opérateurs permettant de passer d'une granularité à une autre n'ont pas été défini dans ce modèle. En ce qui concerne la gestion de l'identité des entités, si sa visualisation et les opérateurs (i.e. naissance, fusion etc.) qui s'y rattachent ne sont pas repris ici, l'identifiant unique associé à une entité permet néanmoins de suivre techniquement (i.e. cet identifiant n'est pas visuel, il fait partie des attributs de l'entité) une entité dans le temps.

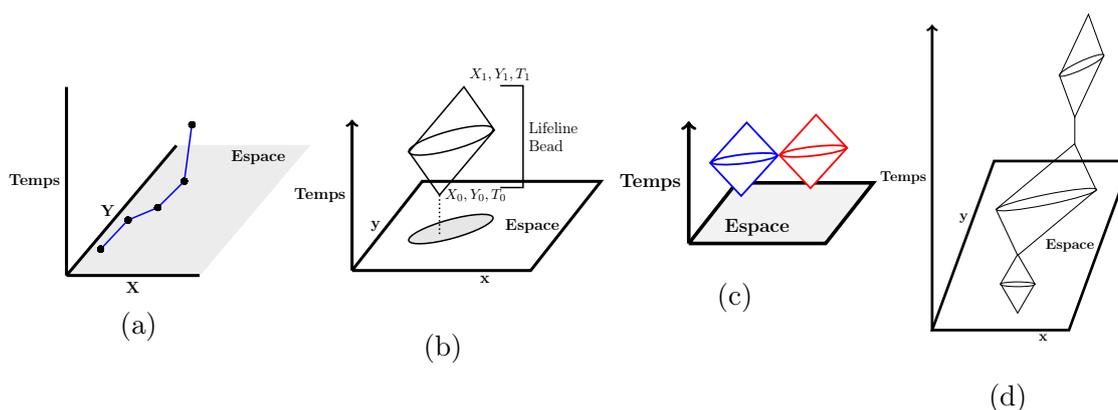


Figure 13 – Modélisation de l'évolution d'entités à différents niveaux de granularité spatio-temporels (Figures adaptées de [Hornsby et Egenhofer, 2002]) : (a) une géospatial lifeline ; (b) un lifeline bead ; (c) deux entités vont potentiellement se rencontrer ; (d) une suite de déplacements spatio-temporels

Ces travaux ([Egenhofer et Al-Taha, 1992, Hornsby et Egenhofer, 1997, Hornsby et Egenhofer, 2000]) ne font que modéliser les changements qui peuvent apparaître

entre (deux) entités. Il est difficile d'en déduire les causes qui sont responsables de ces changements car il est difficile de visualiser un phénomène dans sa globalité, mais seulement selon des temps pris deux à deux. Les modèles de [Hägerstrand, 1970] et [Hornsby et Egenhofer, 2002] sont plus explicites en terme de visualisation globale d'un phénomène, mais ils ne modélisent pas les changements d'états d'une entité, et les relations spatiales et temporelles ne sont pas aisément manipulables. Remarquons tout de même que la plupart de ces modèles (à part celui de [Egenhofer et Al-Taha, 1992]), sont capables, éventuellement en ajoutant les opérateurs ou les requêtes appropriées, de fournir des données quantitatives au niveau global (e.g. statistiques sur l'occupation de l'espace par un attribut en particulier). Finalement, comme l'avait anticipé [Peuquet et Duan, 1995] dès 1995, le tableau de synthèse I.3 montre qu'aucun de ces modèles ne permet d'analyser complètement l'évolution temporelle de l'espace. Un moyen pour essayer de remédier aux problèmes posés par ces modèles est la vision *évènement/processus*.

Modèles :		Snapshot	STC	ST-object	Matrice	CDL	Geolifeline
			Langran 1988 & AI	Worboys 1992	Egenhofer & AI 1992	Hornsby & AI 2000	Hornsby & AI 2002
Représenter	◦ Mouvement	-	-	-	+	--	+
les changements	◦ Chgt. d'état	+	+	++	--	--	+
liés à/au :	◦ l'Identité	NR	-	-	-	++	+
Gestion	◦ Temp.	-	+	--	NR	-	+
granularité :	◦ Spatiale	+	-	+	NR	--	+
Manipuler et analyser		+	-	+	-	-	+
des données ST à							
partir de relations							
temp. et spa.							
Gestion des données :	◦ Initialisation	++	+	-	-	NR	-
	◦ Mise à jour	-	--	-	+	NR	-
	◦ Stockage	-	+	-	-	NR	+

Tableau I.3 – Comparatif des différents modèles basés sur la discrétisation de l'espace.

I.2.2.2 Les modèles basés sur les concepts d'évènements et de processus

[Bestougeff et Ligozat, 1992] définissent un évènement comme un ensemble de processus et un processus comme une action qui modifie une entité. Plus tard, [Galton, 2000] différencie plusieurs types de processus et d'évènements (voir section I.2.1.2).

A partir de ces concepts, [Peuquet et Duan, 1995] proposent le modèle *Event-Based Spatio-temporal Data Model* (ESTDM) basé sur le TRIAD [Peuquet et Wentz, 1994]. Les évènements sont organisés sur une droite temporelle sous la forme d'une liste, et un nouvel évènement est créé si un

changement est détecté. Ce modèle manipule des données de type raster, qui sont liées à chaque évènement (Figure 14). La mise à jour de ce modèle est difficile, et le fait qu'il ne traite que des données raster est assez restrictif. De plus, bien que la représentation d'évènement soit intégrée temporellement, elle est statique et fragmentée (i.e. vision des liens entre évènements difficile).

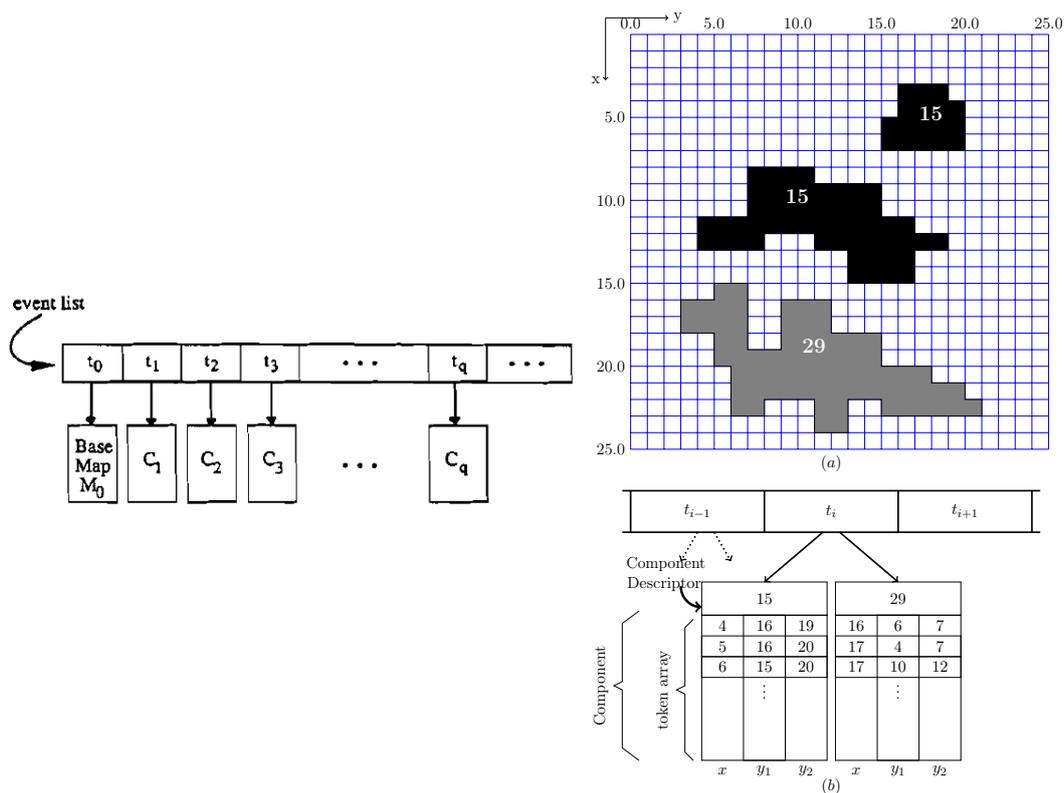


Figure 14 – Modèle *ESTDM* basé sur les évènements et données raster. A gauche la liste des évènements modélisés sous la forme des différents changements qui s’opèrent sur la carte de base ; à droite la façon dont sont représentés ces changements (Figures extraites de [Peuquet et Duan, 1995]).

Or, selon [Claramunt et Thériault, 1995], un modèle de données doit explicitement préserver les liens connus entre les évènements et leurs conséquences et faire clairement la différence entre ce qui est spatial, temporel et thématique. Il faut pouvoir représenter explicitement les processus pour spécifier des relations temporelles entre des entités, et les liens qui existent entre leurs évolutions conjointes. Ils proposent donc une typologie des processus qui s’opèrent sur les entités. Elle comprend 3 classes différentes : déplacement, diffusion et restructuration (Figure 15 a). Le modèle s’appuie sur une structure de temps linéaire et effectue une séparation explicite des dimensions spatiale,

temporelle et thématique. Il préserve aussi les différentes versions⁶ des entités.

[Thériault *et al.*, 1999] étendent ce travail en prenant en compte des ensembles d'entités dans la description des processus, et non les entités de façon individuelle. Ceci les conduit à l'élaboration d'une taxonomie des changements pour ces ensembles d'entités (Figure 15 b), mais l'identification et la qualification des processus décrits dans ces deux approches ([Claramunt et Thériault, 1995, Thériault *et al.*, 1999]) ne sont pas suffisantes. Il faut représenter les entités et les processus liés à leur évolution, en définissant le lien nécessaire pour découvrir les relations de causalités qui produisent les évènements observés. [Hallot et Billen, 2008] et [Muller, 2002] se basent sur un raisonnement qualitatif offrant des constructeurs et des mécanismes de modélisation pour raisonner sur ces changements spatio-temporels, évènements et processus. Dans le même objectif, [Claramunt et Thériault, 1996] ont fait le choix d'un langage pour comprendre et décrire la sémantique des processus composites. Un processus composite est par exemple un feu de forêt qui peut se décomposer en 5 processus : apparition, expansion et déplacement, contraction, disparition. C'est cette décomposition qui est décrite par le langage.

[Worboys, 2005] décrit également un formalisme et un langage basés sur les évènements pour représenter les changements, mais cette fois-ci focalisés dans l'espace. [Jiang et Worboys, 2009] poursuivent ces travaux liés aux changements spatiaux. Ils différencient et spécifient les différents types d'évènements spatiaux qualifiés de *topological event*. Ils identifient un changement spatial de la manière suivante : pour chaque snapshot, une abstraction de la structure de l'évènement est récupérée et modélisée sous la forme d'un arbre. Les modifications spatiales de cet évènement sont caractérisées par les changements (morphismes) entre les arbres (Figure 16). A partir de ces modélisations de changements de bases (séparation, fusion, insertion etc.), il est possible de représenter des évolutions plus complexes en combinant ces arbres.

Ces approches sont plutôt conceptuelles et [Worboys et Hornsby, 2004] font un parallèle entre le concept objet des bases de données en informatique et le concept d'évènement. Mais tous ces modèles basés sur les évènements et les processus ne remplissent pas le cahier des charges que nous nous sommes fixé (Tableau I.4). Ils ne fournissent aucun mécanisme de représentation et d'analyse des réseaux qui caractérisent et relient les évolutions entre entités spatiales.

6. Une version d'une entité est l'entité à un temps donné, il existe autant de versions que de temps auxquels existe l'entité.

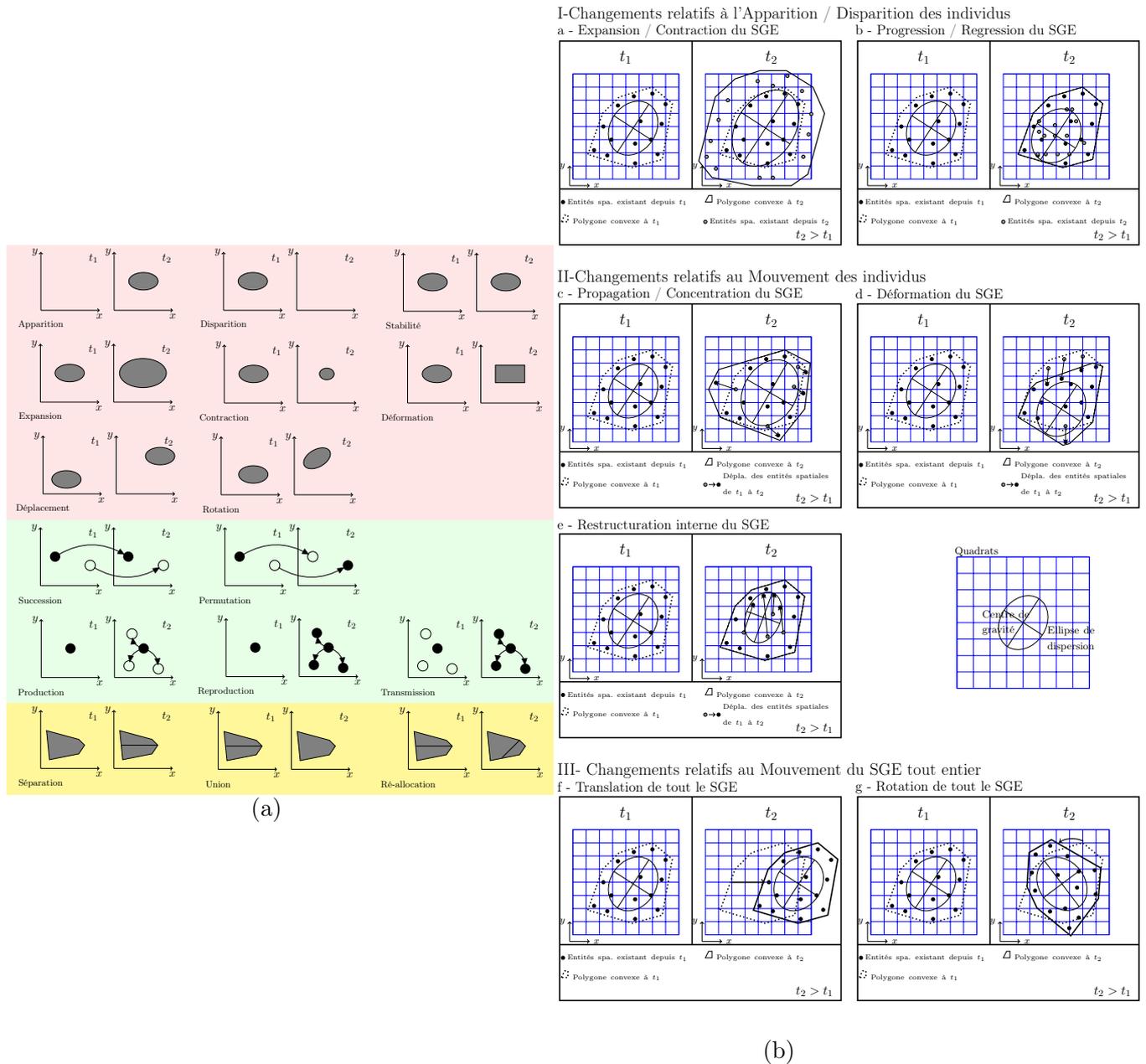


Figure 15 – (a) Typologie des processus pour une entité selon [Claramunt et Thériault, 1995] : rose, évolution d'une entité ; bleu, relations fonctionnelles entre entités ; jaune, évolution de la structure spatiale d'entités ; (b) Taxonomie des processus d'un ensemble d'entités (*Set of Geographical Entities SGE*) selon [Thériault *et al.*, 1999]

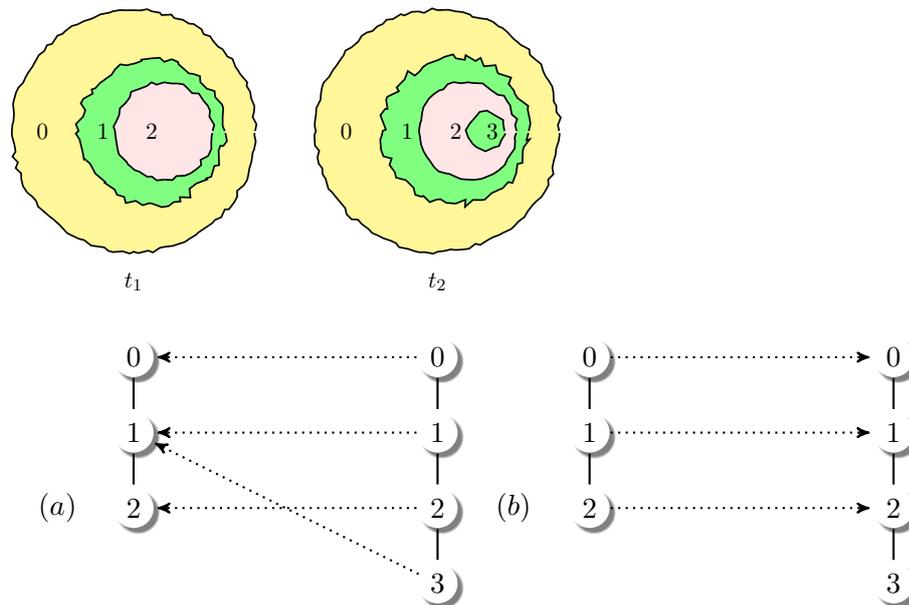


Figure 16 – Modélisation des changements sous la forme de morphismes d'arbres selon [Jiang et Worboys, 2009] : (a) cas d'une séparation (morphisme surjectif) ; (b) cas d'une insertion (morphisme injectif)

Modèles :		ESTDM		Typologie		Taxonomie		Processus composite		Topological event	
		Peuquet & Al 1995	Claramunt & Al 1995	Claramunt & Al 1995	Theriault & Al 1999	Theriault & Al 1999	Claramunt1996 & Al	Claramunt1996 & Al	Worboys & Al 2009	Worboys & Al 2009	
Représenter	o Mouvement	-	++	++	++	++	++	++	++	+	+
les changements	o Chgt. d'état	-	++	++	++	++	++	++	++	-	-
liés à/au	o Identité	--	+	+	+	+	+	+	+	+	+
Gestion	o Temp.	+	-	-	-	-	--	--	--	-	-
granularité :	o Spatiale	--	--	--	--	--	--	--	--	+	+
Manipuler et analyser		+	++	++	++	++	++	++	++	-	-
des données ST à											
partir de relations											
temp. et spa.											
Gestion des données :	o Initialisation	-	+	+	+	+	-	-	-	+	+
	o Mise à jour	-	+	+	+	+	-	-	-	+	+
	o Stockage	-	+	+	+	+	+	+	+	+	+

Tableau I.4 – Comparatif des différents modèles basés sur les concepts d'évènement et de processus.

Modéliser l'évolution d'entités dans l'espace et le temps nécessite de mettre en évidence les changements qui s'opèrent sur ces entités dans ces deux dimensions. L'approche *top down* consiste, à partir d'un phénomène que l'on veut observer, à en déduire les relations et les entités utiles à son analyse. Elle implique de pouvoir définir les liens nécessaires pour établir les causes qui produisent les évènements observés. La description de ces relations et, plus généralement des relations qui existent entre entités spatiales, doit donc être modélisée et exploitée. Un outil permettant d'extraire et de décrire les propriétés sous-jacentes à ces relations est une représentation sous forme de graphe, car cette modélisation peut représenter explicitement les relations entre les entités.

I.3 Les graphes comme support de modélisation spatio-temporelle

L'évolution d'entités spatiales dans le temps génère des réseaux de transition qui peuvent être assimilés à des trajectoires spatio-temporelles [Stefanakis, 2003]. Il paraît naturel de chercher à exploiter ces réseaux qui expriment de nombreuses propriétés encore mal étudiées. La théorie des graphes sous-jacente à ce type de modélisation offre des capacités encore peu exploitées dans les SIG, même si le domaine spécifique de la géographie vue sous forme de réseaux pour étudier les phénomènes dans le temps et l'espace est prometteur [Batty, 2003]. On peut citer en particulier les travaux relatifs au traitement spatio-temporel dans les réseaux sociaux ([Worboys, 2010, Goodchild, 2010]).

Cette modélisation d'un phénomène sous la forme d'un réseau (i.e. graphe) nécessite d'identifier les entités (sommets) et les relations (arêtes) qui les lient dans l'espace et le temps. Par exemple [Jiang *et al.*, 2000] modélisent les espaces urbains avec un graphe dit de *connectivité* où chaque espace est représenté par un sommet, et deux sommets sont reliés si les espaces qui leur sont associés communiquent (Figure 17). Ils exploitent la théorie et les outils de la *syntaxe spatiale*. Introduite par [Hillier et Hanson, 1984]), elle est encore peu utilisée dans les SIG (e.g. [Batty, 2004]). La syntaxe spatiale est basée sur des outils liés à la théorie des graphes (e.g. indice de connectivité) et permet d'analyser l'impact social lié à l'organisation spatiale de bâtiments. En particulier dans cet article de [Jiang *et al.*, 2000], ils font un parallèle entre l'organisation spatiale d'un bâtiment et le déplacement des personnes au sein de ce bâtiment.

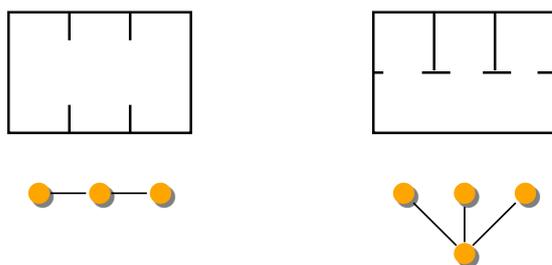


Figure 17 – Graphe de connectivité modélisant des espaces

Un graphe de connectivité analogue est utilisé par [Jiang et Claramunt, 2004a] dans le cadre de l'étude d'un réseau routier. Le graphe est légèrement différent dans le sens où les sommets représentent les rues d'une ville, et une arête est définie entre deux sommets s'il existe une intersection entre les deux rues. Dans cet article, l'intérêt de la théorie des graphes réside dans les opérateurs qui

sont utilisés pour étudier les propriétés structurelles du graphe, et donc du réseau routier sous-jacent. Par exemple, les auteurs calculent la connectivité des sommets : plus un nœud possède de connections et plus il est important du point de vue de la structure du réseau (e.g. conséquences d'un carrefour bloqué). [Jiang et Claramunt, 2004b] poursuivent la démarche mais utilisent un autre concept propre à la théorie des graphes, à savoir les réseaux de type *small world* et *scale free* qui ont des caractéristiques structurelles particulières. Par exemple dans le cas du réseau de type *small world*, un nœud ne possède pas beaucoup de voisins, mais beaucoup de sommets sont accessibles depuis ce nœud avec une distance (i.e. un nombre d'arêtes) faible. Connaître ces propriétés peut s'avérer très intéressant lors de la construction d'un réseau routier.

Dans tous ces exemples ([Jiang et al., 2000, Jiang et Claramunt, 2004a, Jiang et Claramunt, 2004b]), c'est essentiellement de la structure spatiale dont il est question. La prise en compte de la dimension temporelle implique de définir différemment le graphe modélisant le phénomène.

[Spéry et al., 2001] créent des graphes orientés pour suivre des *filiations* d'entités sur un cas d'étude cadastrale. Par filiation, on entend ici filiation spatiale. C'est-à-dire que l'on se base sur des relations purement spatiales pour dire si une entité a une relation de filiation avec une autre, et non pas sur le concept d'identité défini précédemment (même si on conserve tout de même un identifiant unique par entité). Les sommets du graphe représentent les entités (i.e. les parcelles), et deux sommets consécutifs (temporellement) sont reliés par une arête si l'espace de la première parcelle est en relation spatiale avec l'espace de la seconde (Figure 18). Comme il s'agit de parcelles, la filiation est essentiellement spatiale, et la dimension temporelle est clairement modélisée par l'orientation des arêtes. Ce modèle permet d'identifier des changements qu'ils qualifient d'élémentaires (e.g. division) ou complexes (e.g. ré-allocation), mais il est difficile de représenter un déplacement d'entité. Par ailleurs, la relation qui existe entre les entités est imposée par le domaine d'application (i.e. ici seulement spatiale), et le modèle n'intègre pas de relation entre entités présentes à un même temps, ni de relations entre des temps non consécutifs.

[Sriti et al., 2005] intègrent également de façon explicite la dimension temporelle avec un graphe orienté. Dans ce but, ils proposent de combiner plusieurs graphes, respectivement les relations spatiales entre entités à un instant donné, et les relations de filiations (spatiales) entre entités entre des temps consécutifs. Pour chaque temps, le graphe de relations spatiales est décrit avec les parcelles

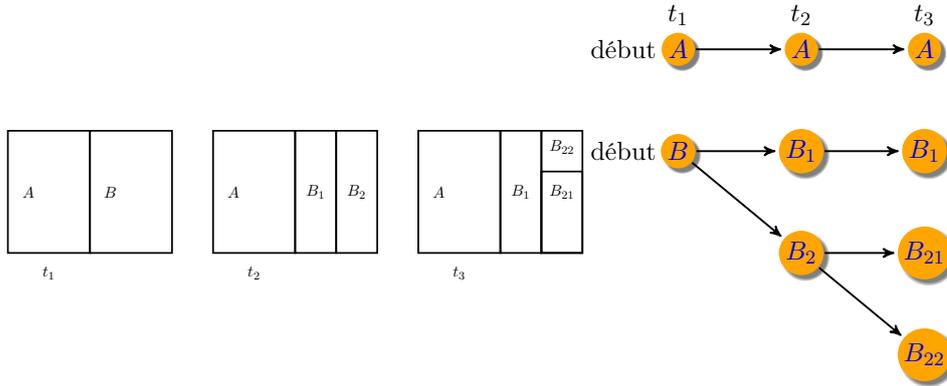


Figure 18 – Graphe orienté modélisant l'évolution de parcelle selon une filiation spatiale

comme sommets, et deux sommets sont reliés entre-eux si les parcelles se touchent. La combinaison de ce graphe et du graphe de filiations permet de décrire qualitativement la structure du graphe résultant (e.g. opérateurs de théorie des graphes comme le calcul de la centralité). Dans ce travail, les filiations sont limitées au domaine spatial, et l'exploitation du graphe ne permet pas d'analyser les processus responsables de l'évolution modélisée.

Le modèle qui nous a particulièrement intéressé compte tenu de nos objectifs (i.e. approche entité, relations spatiales et d'identité dans le temps) est le modèle de [Stell, 2003]. Il formalise ce type d'approche spatio-temporelle qui utilise les graphes, en proposant un modèle générique de type snapshot. Le modèle de temps proposé peut ne pas se restreindre à la structure linéaire (i.e. on peut théoriquement avoir des domaines de temps en parallèle). Les entités (sommets) sont reliées entre elles au travers du temps (pas forcément entre temps consécutifs) si elles partagent une certaine relation (non clairement définie, spatiale, filiation familiale etc.) qu'il appelle relation *support* (Figure 19). Il en découle une formalisation des processus de base d'une évolution (e.g. naissance, séparation, fusion, mort etc.) et la possibilité de modéliser des phénomènes variés au vu de la généralité de cette représentation. Cependant il n'est pas possible de combiner plusieurs type de relations au sein d'un même graphe (e.g. spatiale et filiation).

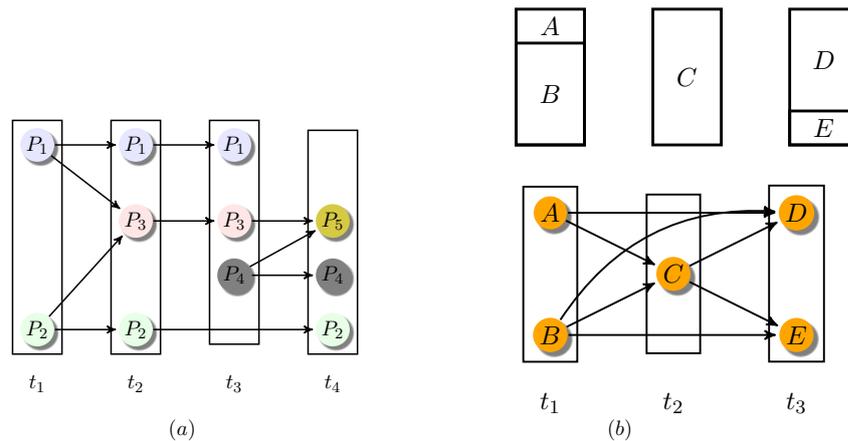


Figure 19 – Modélisation de l'évolution d'entité selon [Stell, 2003] : (a) Evolution d'une famille d'individus, la relation support représente les liens génétiques, P_1 et P_2 sont les parents de P_3 ; Il n'y a pas d'information concernant P_2 en t_3 ; (b) Evolution de parcelles, la relation support représente la superposition (OV), entre autres on a $C OV A$, $C OV B$ et $D OVC$

Comme précédemment, il est possible de synthétiser ces approches dans un tableau récapitulatif (Tableau I.5) basé sur la définition de caractéristiques nécessaires à un modèle spatio-temporel. Il en ressort qu'aucun de ces modèles basés sur la théorie des graphes ne peut représenter le mouvement d'une entité. Par ailleurs, la manipulation de données spatio-temporelles par le biais de relations spatiales et temporelles reste limitée, et ne permet pas l'analyse des processus responsables de l'évolution du phénomène modélisé.

Modèles :		Graphes de connectivité		Graphe ST de filiations	Graphe ST (1)	Graphe ST (2)
		Jiang & Al 2000 et Jiang & Al 2004		Spery & Al 2001	Striti & Al 2005	Stell 2003
Représenter	o Mouvement	-		-	+	+
les changements	o Chgt.d'état	+		+	+	+
liés à/au	o l'Identité	NR		++	+	++
Gestion	o Temp.	-		-	-	+
granularité :	o Spatiale	-		-	-	-
Manipuler et analyser des données ST à partir de relations temp. et spa.		+		+	+	+
Gestion des données :	o Initialisation	+		+	+	+
	o Mise à jour	+		+	-	-
	o Stockage	+		+	+	+

Tableau I.5 – Comparatif des différents modèles basés sur les concepts de théorie des graphes. Les trois derniers modèles ne représentent pas explicitement un mouvement mais sont capables d'extraire un changement spatial entre deux temps.

L'approche spatio-temporelle modélisée au travers de la théorie des graphes offre des capacités d'études des phénomènes représentés. En effet, la modélisation par les graphes, basée sur l'approche entités/rerelations, représente les relations spatiales et temporelles et propose de nombreux outils d'analyse rigoureusement définis (e.g. opérateur de centralité, petits mondes). L'évolution d'entités est généralement vue comme un changement dans la structure spatiale (e.g. géométrie, déplacement, connectivité) de ces entités. Cependant il n'est pas pertinent de dissocier les changements spatiaux de l'identité des entités associées au phénomène. Nous voudrions pouvoir intégrer ces deux concepts simultanément pour suivre ces entités dans le temps. Le modèle de [Stell, 2003] propose une solide formalisation et le caractère générique de la relation 'support' qu'il définit, nous semblent propice à l'élaboration d'un modèle plus complexe. Celui-ci doit nous permettre de modéliser et d'étudier l'évolution d'entités, dans l'espace, le temps, et de conserver leur filiation du point de vue de leur identité.

Un graphe spatio-temporel pour modéliser l'évolution d'entités

Sommaire

II.1	Spécification du modèle	36
II.1.1	Relations de l'espace et du temps	38
II.1.1.1	Définition du domaine temporel	38
II.1.1.2	Relation de connexion spatiale	38
II.1.1.3	Relation spatio-temporelle	39
II.1.1.4	Relation de filiation	40
II.1.2	Le graphe spatio-temporel	42
II.1.3	Un exemple de modélisation : propagation d'une population de ronces	44
II.2	Fonctions du modèle	47
II.2.1	Fonctions basées sur les relations	48
II.2.1.1	Définition générale	48
II.2.1.2	Fonctions basées sur la relation spatiale de connexion	49
II.2.1.3	Fonctions basées sur la relation spatio-temporelle	50
II.2.1.4	Fonctions basées sur la relation de filiation	51
II.2.1.5	Application des fonctions basées sur les relations à l'exemple sur l'évolution de la population de ronces	52
II.2.2	Fonctions d'identification des processus	55
II.2.3	Fonction de jointure	56
II.2.3.1	Fonctions d'union et d'intersection	57
II.2.3.2	Fonction de jointure conditionnelle	57
II.2.3.3	Un exemple d'application de la jointure	59

II.2.4	Routes et relations	64
II.2.4.1	Définition générale et routes spatio-temporelles	64
II.2.4.2	Exemple d'utilisation des routes	68

Le chapitre précédent a montré que l'espace et le temps étaient indissociables et nécessaires à l'analyse de l'évolution d'entités spatio-temporelles. Les changements qui s'opèrent sur ces entités sont caractérisés par des modifications spatiales (e.g. liées à la position, à l'empreinte spatiale) ou liées à leur identité. Les différents modèles décrits précédemment ne gèrent pas de façon simultanée ces relations. Nous proposons un modèle qui intègre la totalité et distingue les relations liées à l'espace et à l'identité, d'une part dans le temps (relations spatio-temporelles, filiations temporelles) et d'autre part à un instant donné (relations spatiales, filiations). Ce modèle est basé sur certains concepts formels proposé par [Stell, 2003] que nous commençons par détailler au début de ce chapitre (l'introduction de la section II.1). Ensuite, nous définissons notre modèle. La dernière partie de ce chapitre décrit les différents types de fonctions qui peuvent être appliquées en vue de l'analyse du phénomène modélisé.

II.1 Spécification du modèle

Le modèle proposé par [Stell, 2003] est défini sur un domaine de temps (T, \leq) avec $T = \{t_i, \dots, t_j\}$ où $i < j$, et \leq l'ordre partiel sur T . Cet ordre partiel implique que l'on peut théoriquement utiliser des alternatives à la représentation linéaire du temps, avec en particulier la structure du temps parallèle (Figure 20).

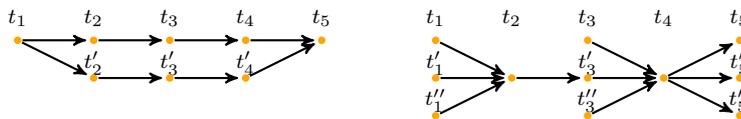


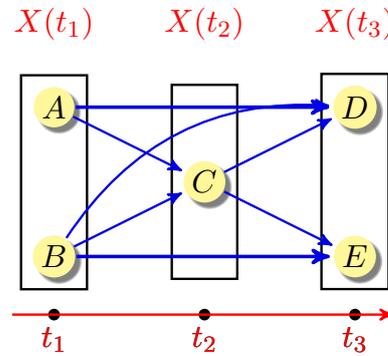
Figure 20 – Le domaine de temps défini pour le modèle de [Stell, 2003] autorise d'autres représentations du temps que la structure linéaire classique. A gauche on voit une partie du domaine de temps qui évolue en parallèle, à droite les évolutions se rejoignent et se séparent.

Sur ce domaine de temps, [Stell, 2003] définit un *dynamic set* qui représente un ensemble d'entités évoluant au cours du temps. Il comprend une fonction $f : \sigma \rightarrow T$ qui renvoie pour chaque entité dans σ le temps dans T où elle existe, ainsi qu'un domaine de temps (T, \leq) associé. L'ensemble des entités existant à un temps $t_i \in T$ est noté $X(t_i)$.

Les entités sont reliées entre les temps au moyen de la relation dite "*support*". Si cette relation

existe entre deux entités a et b présentes respectivement au temps t_i et t_j , $i < j$, cela signifie que a "supporte" b c'est-à-dire que b dépend d'une certaine manière de l'existence de a . Cette relation est très clairement générique et son interprétation dépend de l'application. Néanmoins, [Stell, 2003] la pose comme étant transitive. En ce qui concerne la représentation, bien que les arêtes relatives à cette relation ne soient pas orientées, la relation suit la même orientation que celle du temps (de la gauche vers la droite du graphe).

L'ensemble des concepts liés à ce modèle sont résumés par la Figure 21. Il nous fournit les bases nécessaires pour définir notre propre modèle, en particulier concernant la déclinaison de la relation *support* en relation spatiale, spatio-temporelle et de filiation.



T	Le domaine de temps (T, \leq) :	défini par l'ensemble discret fini de temps T et associé à l'ordre partiel \leq .
X	$\forall t, X(t)$:	l'ensemble des entités au temps t
S	La relation support :	A est le support de C si l'existence de C dépend de celle de A .
B	La relation support $X(t_i, t_j)$:	relation binaire définie entre deux temps $t_i \leq t_j$
P	La relation support transitive :	$\forall i < j < k$, soient $x \in X(t_i), x' \in X(t_j), x'' \in X(t_k)$ $X(t_i, t_j)$ et $X(t_j, t_k) \Rightarrow X(t_i, t_k)$
f	La fonction $f : \sigma \rightarrow T$:	renvoie pour chaque entité le temps où elle existe, elle implique qu'une entité ne peut exister qu'en un seul exemplaire à chaque temps
D	Le modèle (Dynamic Set) :	défini comme l'association du domaine de temps (T, \leq) et de la fonction f .

Figure 21 – Résumé de la formalisation du modèle spatio-temporel de [Stell, 2003] : chaque lettre majuscule correspond à un concept que nous avons réutilisé tel quel ou adapté dans notre modèle. Cette notation est reprise dans le texte.

II.1.1 Relations de l'espace et du temps

II.1.1.1 Définition du domaine temporel

Soit l'ensemble des temps $T = \{t_1, \dots, t_n\}$ avec $n \in \mathbb{N}$. T est fini et, sans faire de généralités, on pose pour tout $i \in [1, n]$ les points t_i, t_{i+1}, \dots, t_n consécutifs, et $t_i < t_{i+1}$. Ce domaine de temps est équivalent à celui défini dans le modèle de [Stell, 2003] (concept **T** de la Figure 21), mais pour ce qui nous concerne, nous ne traiterons que d'une représentation linéaire du temps.

Alors que la relation support (concept **S**) peut être utilisée de différentes façons sans distinction (e.g. relation spatiale, liée à l'identité), nous allons la spécialiser et lui permettre de revêtir plusieurs types de relations (e.g. spatiale, liée à l'identité), et faire coexister ces relations.

Remarque : Lors de l'écriture d'une relation, s'il n'y a pas d'ambiguïté on peut omettre l'indication temporelle t_i . Par ailleurs, les relations définies dans le temps (spatio-temporelles et filiations temporelles) ne sont pas nécessairement définies entre deux temps consécutifs.

II.1.1.2 Relation de connexion spatiale

Nous avons choisi de supposer que chaque entité est associée à une région de \mathcal{R}^2 , mais notre modèle ne l'impose pas a priori. Il existe entre ces entités une relation de connexion C dont la sémantique dépend de l'application, mais C sera a priori toujours symétrique et réflexive. Dans une première approche C peut être spécifiée comme relation du *RCC8* [Randell *et al.*, 1992] qui comprend sept cas de connexion et un cas de déconnexion DC (Table 22). Dans ce cas on pose $\rho_s = C$ définie sur les sept premiers cas. La négation de cette relation, ϕ_s , correspond au dernier cas DC . On pose Γ l'ensemble de ces relations (ici le *RCC8*) valides pour la relation spatiale.

Soit $x, y \in \sigma$ deux entités :

$x DC y$	x déconnecté de y
$x EC y$	x connecté par frontière à y
$x TPP y$	x contenu dans y et ayant une frontière tangente à y
$x NTPP y$	x contenu dans y et n'ayant pas de frontière tangente à y
$x PO y$	x chevauche y
$x EQ y$	x égal y
$x TPP_i y$	$x TPP^{-1} y$, y contenu dans x et ayant une frontière tangente à x
$x NTPP_i y$	$x NTPP^{-1} y$, y contenu dans x et n'ayant pas de frontière tangente à x

Figure 22 – Les relations du *RCC8*

Soit σ l'ensemble de toutes les entités présentes sur T . Si deux entités $a, b \in \sigma$ telles que a et b existent au même instant $t \in T$, et si elles sont en relation de connexion spatiale ρ_s , alors on écrit :

$a_t \rho_s b_t$ **et** $a_t \not\rho_s b_t$ **sinon**

L'arête entre les deux entités Figure 23 peut être étiquetée avec la relation topologique souhaitée.



Figure 23 – Représentation de ρ_s à un temps t : à gauche une relation de connexion spatiale entre a et b ; à droite il n'y a pas de relation de connexion spatiale entre ces entités.

II.1.1.3 Relation spatio-temporelle

La relation spatio-temporelle repose sur la même base de définition que la relation spatiale, et donc sur le même ensemble de relations Γ , mais entre des entités définies à des temps différents. La distinction entre la relation spatiale et la relation temporelle permet, lors de la spécification du modèle, d'être plus précis et de pouvoir les utiliser sans ambiguïté. En effet, dans le modèle il existe une distinction importante entre la relation spatiale qui capture la façon dont les entités sont en relation à un temps donné, et la relation spatio-temporelle qui exprime comment des entités sont en relation à des temps différents. En particulier, la relation spatio-temporelle n'est pas contrainte d'utiliser le même type de connexion que la relation spatiale (même si dans le cadre général de nos définitions nous avons pris le même¹ ensemble de relations Γ). Par exemple, lors de l'évolution d'une population de dunes marines, deux dunes ne peuvent pas se chevaucher à un instant donné, par contre c'est tout à fait possible entre des temps différents.

Soient les entités spatiales a et b existant respectivement au temps t_i et t_j avec $t_i < t_j$. Si l'espace associé à l'entité spatiale b est en relation de connexion spatiale ρ_s avec celui de a alors il existe une relation spatio-temporelle entre ces deux entités (Figure 24 et 25) et on écrit :

$a_{t_i} \rho_{st} b_{t_j}$ **et** $a_{t_i} \not\rho_{st} b_{t_j}$ **sinon.**

1. Ce choix est cohérent dans le sens où baser les relations spatiales et spatio-temporelles sur l'ensemble Γ , n'impose pas a priori d'utiliser toutes les relations de Γ .

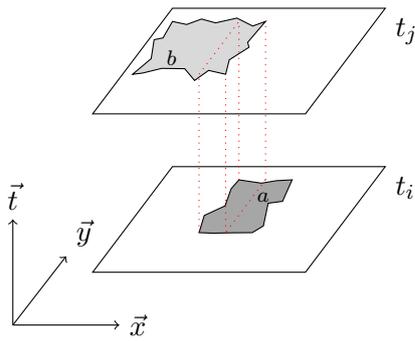


Figure 24 – L'espace occupé par a au temps t_i est en relation de connexion spatiale ρ_s avec celui occupé par b au temps t_j , il existe donc une relation spatio-temporelle ρ_{st} entre ces deux entités.



Figure 25 – Représentation de la relation spatio-temporelle entre deux entités : à gauche l'entité a au temps t_i a une relation spatio-temporelle avec l'entité b au temps t_j , à droite a et b n'ont pas de relation spatio-temporelle. Comme pour la relation spatiale, l'arête entre les deux entités en relations peut être étiquetée avec la relation topologique souhaitée.

II.1.1.4 Relation de filiation

De manière générale, la relation de filiation est définie sur le concept d'identité. La relation d'identité entre deux entités peut être vue comme une relation de dépendance qui est la transmission de la parenté lorsqu'une entité descend d'une autre. Cette notion de descendance est à préciser suivant le domaine d'application. La relation de filiation est définie à un instant ou dans le temps.

Soient les entités spatiales a et b existant respectivement au temps t_i et t_j avec $t_i \leq t_j$ (Figure 26), si l'entité b dépend de l'existence de a (si b descend de a) alors a et b appartiennent à la même filiation et on écrit : $a_{t_i} \rho_f b_{t_j}$ et $a_{t_i} \rho_f^/ b_{t_j}$ sinon.

On différencie deux types de filiations, la continuation et la dérivation (Figure 27) d'où $\rho_f \in \Phi = \{\gamma, \delta\}$ tel que :

- **La filiation de type continuation γ** : $a_{t_i} \gamma b_{t_j}$ signifie que a et b entretiennent une relation de filiation de type continuation, c'est-à-dire que a et b ont la même identité, il s'agit de la même entité mais à des temps différents. Nous considérons que ce type de filiation n'est défini que si $t_i < t_j$, c'est à dire que concrètement deux entités ayant la même identité ne peuvent coexister à un même temps. Au-delà des considérations philosophiques à l'origine d'un tel choix (e.g. deux vrais jumeaux = deux identités différentes ?), c'est également dans un souci de pouvoir identifier chaque entité à chaque instant que nous avons pris cette décision. De plus nous restons en accord avec le fait qu'il ne peut exister qu'un exemplaire de chaque entité à chaque temps qu'implique I .

- **La filiation de type dérivation δ** : $a_{t_i} \delta b_{t_j}$ signifie que a et b entretiennent une relation de filiation de type dérivation, b est une dérivation de a , c'est-à-dire que b dérive de a (et potentiellement d'une autre entité s'il s'agit de relation parents-enfant). On considère qu'une partie de l'identité de a est présente dans b .

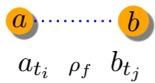


Figure 26 – Représentation de ρ_f : à gauche, une entité a a une relation de filiation avec l'entité b ; sur la droite, il n'y a pas de relation de filiation entre ces entités.

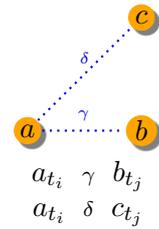


Figure 27 – Relations de filiations entre deux entités : continuation entre a et b , dérivation entre a et c .

Afin de spécifier davantage cette relation de filiation, nous posons deux contraintes. Elles sont assez générales pour pouvoir s'appliquer à un grand nombre d'applications, et nous permettent ainsi de mieux cadrer la relation (Figure 28) :

- Une entité ne peut pas être en relation de continuation dans le temps avec plus d'une seule entité. A défaut, cela autoriserait deux entités de même identité à un même instant, ce que nous avons posé comme impossible (Figure 28a).
- Une entité issue d'une relation de continuation ne peut être issue d'une autre relation de filiation. Aucune autre entité ne peut venir l'enrichir sinon son identité en serait affectée², et il ne s'agirait plus d'une continuation mais d'une dérivation (Figure 28b).

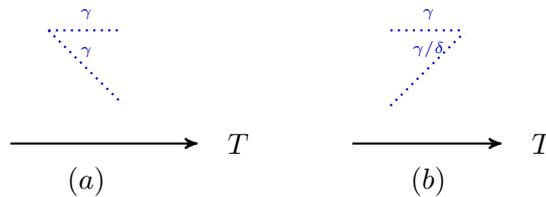


Figure 28 – Relations non admises par le modèle : (a) Cas de la division lors d'une relation de continuation ; (b) Cas de la fusion lors d'une relation de continuation

2. Cette interprétation nous est propre, nous avons bien précisé que ces contraintes sont a priori arbitraires. En particulier pour cette deuxième contrainte, dans le cas de l'absorption (e.g. un individu qui mange une pomme), il est tout à fait discutable que l'identité de l'individu soit réellement différente. Nous avons essayé ici de rester le plus générique possible ; concrètement, il faudrait préciser la dérivation et traiter l'absorption comme une dérivation particulière.

II.1.2 Le graphe spatio-temporel

Nous avons posé les définitions permettant d'établir notre modèle. Il reste à intégrer ces définitions dans un concept global de *spatio-temporal set* relatif au graphe spatio-temporel.

A partir des concepts³ **D**, **X** et **B**, on définit un *spatio-temporal set* comme étant un *dynamic set* X sur un domaine de temps T tel que, pour σ l'ensemble de toutes les entités sur T , $X = \sigma$ et $X(t)$ soit égal à l'ensemble des entités présentes à l'instant t . Pour tous $t, t' \in T$, $t \leq t'$, on pose les relations $X(t, t')$ telles que $X(t, t') = \rho(t, t') = r | r \in \{\rho_s, \rho_f, \rho_{st}\}$. On définit sur ces concepts liés au *spatio-temporal set*, un graphe spatio-temporel $G_{ST} = (\sigma, E_{ST}, T)$ avec $\forall x, y \in \sigma$, $E_{ST} = \{(x, y) | x \rho y\}$.

Les relations dans le modèle et les sous graphes associés

- **La relation spatiale** : La relation de connexion spatiale est spécifiée comme suit dans le modèle, avec pour chaque $X(t)$, $t \in T$, $a, b \in X(t)$ et on a $a_t \rho_s b_t$ ou $a_t \not\rho_s b_t$.

Le sous graphe associé est $G_s = (\sigma, E_s, T)$ avec $\forall x, y \in \sigma$, $E_s = \{(x, y) | x \rho_s y\}$.

- **La relation spatio-temporelle** : La relation spatio-temporelle est spécifiée comme suit, $\forall t < t'$, entre deux entités $a \in X(t)$ et $b \in X(t')$, avec $a_t \rho_{st} b_{t'}$ ou $a_t \not\rho_{st} b_{t'}$.

Le sous graphe associé est $G_{st} = (\sigma, E_{st}, T)$ avec $\forall x, y \in \sigma$, $E_{st} = \{(x, y) | x \rho_{st} y\}$.

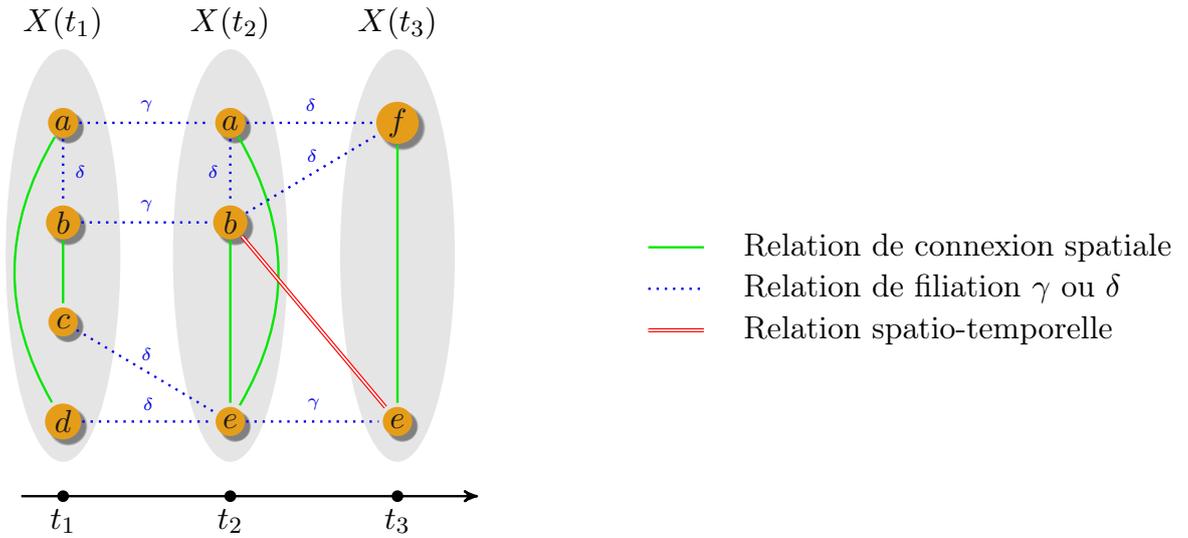
- **La relation de filiation** : la relation temporelle de filiation est spécifiée comme suit, entre deux entités $a \in X(t)$ et $b \in X(t')$, avec $a_t \rho_f b_{t'}$, $\rho_f \in \Phi$, ou $a_t \not\rho_f b_{t'}$.

Le sous graphe associé est $G_f = (\sigma, E_f, T)$ avec $\forall x, y \in \sigma$, $E_f = \{(x, y) | x \rho_f y\}$.

On note que $E_{ST} = E_s \cup E_{st} \cup E_f$. Un exemple d'un tel graphe G_{ST} reprenant les relations définies ci-dessus est proposé Figure 29.

A propos de la transitivité des relations Contrairement au concept **P** où la relation support doit être transitive, nous permettons que les relations utilisées dans le modèle ne le soient pas forcément. La relation spatiale (et donc spatio-temporelle) n'est pas toujours transitive comme l'illustre la Figure 30. Sur la Figure 30a, la relation spatiale est transitive : L'entité e_1 à t_1 est en relation spatio-temporelle avec e_2 à t_2 , celle-ci est en relation spatio-temporelle avec l'entité e_3 à t_3 , et les entités e_1 et e_3 sont également en relation spatio-temporelle. C'est ce qui diffère de la Figure 30b, où l'entité e_1 n'a pas de relation spatio-temporelle avec l'entité e_3 , il n'y a donc pas transitivité

3. **D** : *Dynamic set*, association du domaine de temps (T, \leq) et de la fonction $f : \sigma \rightarrow T$; **X** : $\forall t, X(t)$, l'ensemble des entités au temps t ; **B** : La relation support $X(t_i, t_j)$ relation binaire définie entre deux temps $t_i \leq t_j$



(On peut étiqueter les arêtes spatiales et spatio-temporelles avec les relations du *RCC8*)

Figure 29 – Un exemple de graphe spatio-temporel G_{ST} . Les relations spatiales sont en lignes pleines ; les filiations en pointillés avec les étiquettes γ ou δ pour spécifier s’il s’agit respectivement de continuations ou de dérivations, les relations spatio-temporelles sont en lignes doubles.

de la relation dans ce cas.

Par ailleurs, la relation de filiation est a priori transitive mais il ne faut prendre en considération que des filiations de même type (soit continuation, soit dérivation).

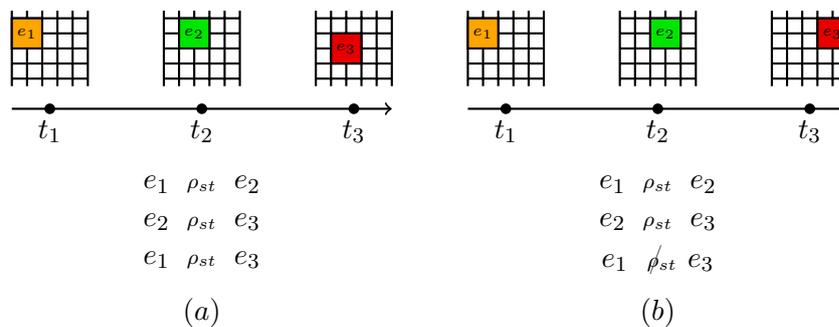


Figure 30 – (a) cas de transitivité de la relation spatiale ; (b) cas de non transitivité de la relation spatiale

De manière générale il faut être conscient que toutes les relations qui existent dans la réalité ne sont pas forcément représentées. L’enjeu est de représenter l’ensemble suffisant et minimal des relations relatives à un phénomène (i.e. en particulier les liens directs). Par exemple dans le cas d’une relation spatiale transitive (Figure 30 a), si la relation spatio-temporelle est déjà représentée entre l’entité e_1 et l’entité e_2 d’une part, et entre l’entité e_2 et l’entité e_3 d’autre part, il n’est pas

nécessaire a priori de matérialiser la relation spatio-temporelle entre l'entité e_1 et l'entité e_3 . Ceci est d'autant plus vrai avec les relations de filiations où la matérialisation de toutes les relations n'est pas nécessaire car l'on retrouve ces liens grâce à des opérateurs (Voir la section II.2). Dans ce cas, faire figurer toutes les relations ferait très rapidement exploser le nombre d'arêtes du graphe.

En revanche, il est préférable d'être exhaustif si l'on sait que le phénomène modélisé implique à la fois des relations transitives et non transitives, ou si l'on souhaite un accès plus direct à l'information (i.e. ne pas représenter une relation impose de passer par l'application d'un opérateur pour la vérifier).

II.1.3 Un exemple de modélisation : propagation d'une population de ronces

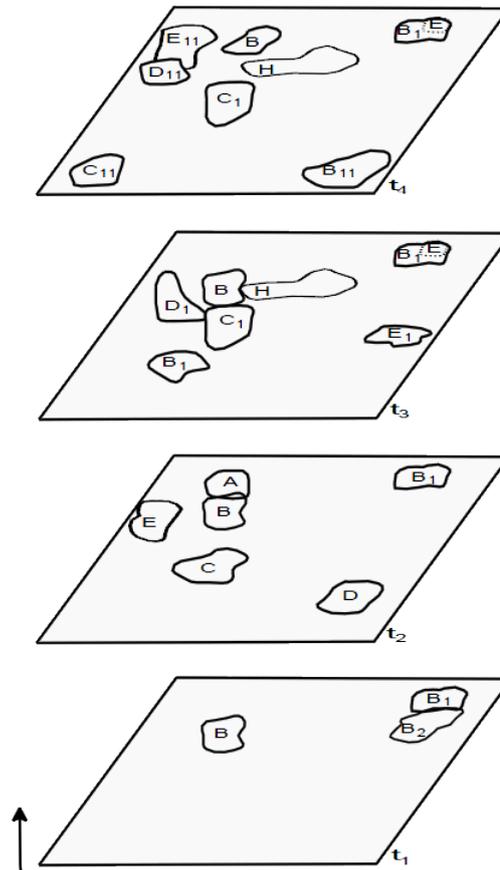
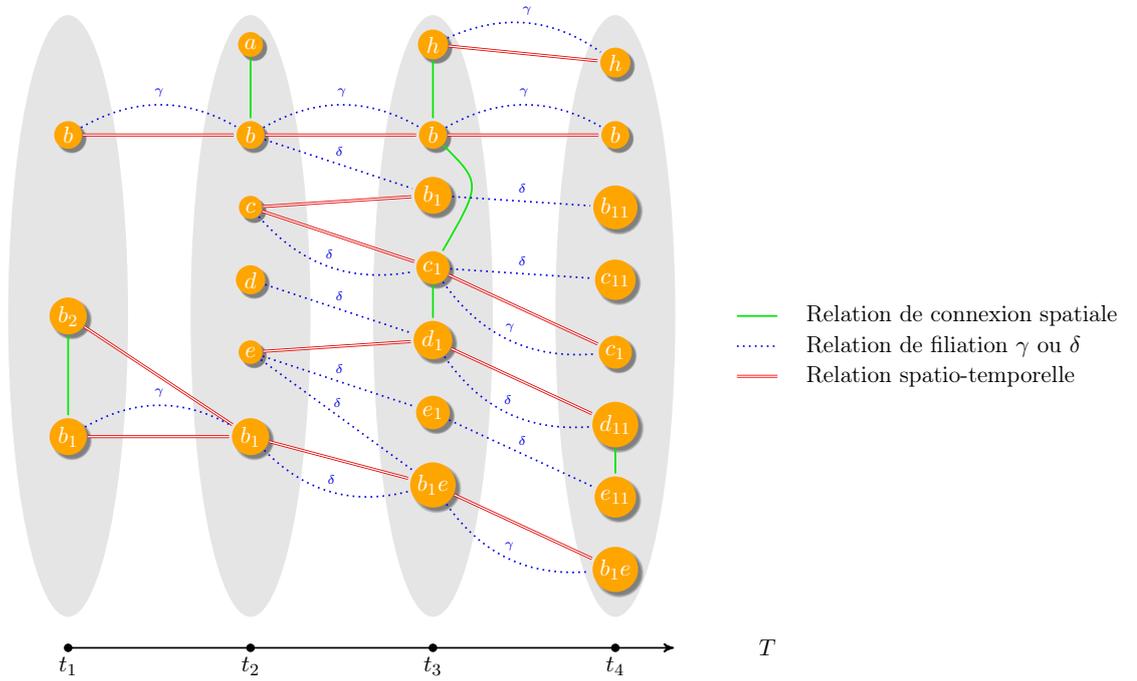


Figure 31 – Evolution d'une population de ronces

A titre d'illustration, nous proposons un exemple⁴ afin d'expliciter ce modèle. Il décrit la

4. Nous nous basons sur cet exemple pour appliquer certains opérateurs de la section II.2.

modélisation de l'évolution d'une population de *ronces communes* observée à quatre temps consécutifs donc $T = \{t_1, \dots, t_4\}$ (Figure 31). Le graphe spatio-temporel associé est présenté Figure 32. Chaque ronce est représentée dans le graphe par un noeud, et les relations spatiales et de filiations entre les ronces sont modélisées par les arêtes du graphe. L'empreinte spatiale d'une ronce est la projection de la plante sur le sol en prenant en compte à la fois les racines souterraines et les branches aériennes. Les ronces possèdent quatre modes de propagation spatiale possibles ([Wehrle, 1985]), dont trois cas de division et un cas de fusion (Figure 33) : le marcottage, le drageonnage, la dispersion des graines et la greffe.



(On peut étiqueter les arêtes spatiales et spatio-temporelles avec les relations du *RCCS*)

Figure 32 – Graphe spatio-temporel G_{ST} associé à l'évolution de la population de ronces.

Nous avons modélisé ces processus par le biais de notre modèle. On suppose deux temps consécutifs $t_j, t_{j+1} \in T$. On appelle x et y des ronces présentes à t_j , et $\forall i \in \mathbb{N}$, il peut exister plusieurs ronces x_i au temps $j+1$ générées par x . x_i est un *surgeon* qui est une plante génétiquement identique⁵ à son parent x :

- **Drageonnage** (Figure 33a) : Il y a propagation par drageonnage si une racine de la plante parente x sort du sol et se détache de x , pour créer une nouvelle ronce x_i .

5. Selon [Wehrle, 1985], les plantes filles sont génétiquement identiques à leur(s) parent(s), préservant ainsi la relation de filiation

Formellement ceci revient à (Figure 34 a) : $x \delta x_i$ et $x \rho_{st} x_i$.

- **Marcottage** (Figure 33b) : Il y a propagation par marcottage si une branche de la ronce parente x se plante dans le sol et se détache de x (i.e. concrètement la branche impliquée meurt), pour créer une nouvelle ronce x_i .

Formellement ceci revient à (Figure 34 a) : $x \delta x_i$ et $x \rho_{st} x_i$.

- **Dispersion des graines** (Figure 33c) : Le fruit d'une ronce x est constitué d'un nombre variable de graines agglomérées. Lorsque ces fruits sont mangés par des animaux (e.g. oiseaux, cerf), leurs excréments peuvent donner naissance à de nouvelles plantes x_i , qui sont donc dérivées⁶ de la plante mère x .

Formellement ceci revient à (Figure 34 b) : $x \delta x_i$.

- **Greffe** (Figure 33d) : Une greffe s'effectue entre deux ronces x et y d'identité différentes. Il y a propagation par greffe si une partie de la ronce y est greffée sur une branche de la ronce x . Il y a alors fusion entre les deux ronces pour créer la plante xy au temps t_{j+1} .

Formellement ceci revient à (Figure 34 c) : $x \delta xy$ et $y \delta xy$. On considère également que la plante greffée reste au même endroit et donc $x \rho_{st} xy$.

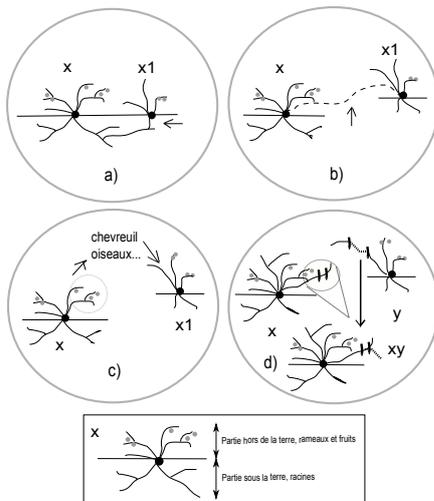


Figure 33 – Vue schématique des mécanismes de propagation des ronces : (a) drageonnage; (b) marcottage; (c) dispersion des graines; (d) greffe

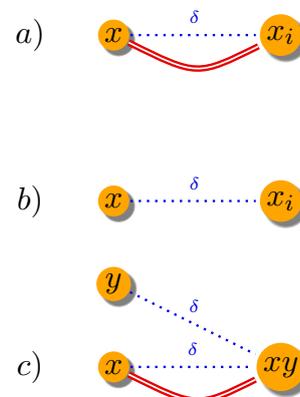


Figure 34 – Modélisation des quatre modes de propagations des ronces (a) drageonnage ou marcottage; (b) dispersion des graines; (c) greffe

Finalement, on identifie les différents processus spécifiés dans l'évolution de la population de ronces (Figure 35). Ceci démontre que l'on peut a priori identifier l'information liée à des processus spécifiques suivant l'application, mais nous n'avons pas ici les outils pour extraire ces processus. Ceci fera l'objet de la partie suivante où nous allons définir ces outils au travers de fonctions.

6. A priori rien n'impose que $x \rho_{st} x_i$, et nous choisirons donc de considérer le cas général où ce n'est pas le cas.

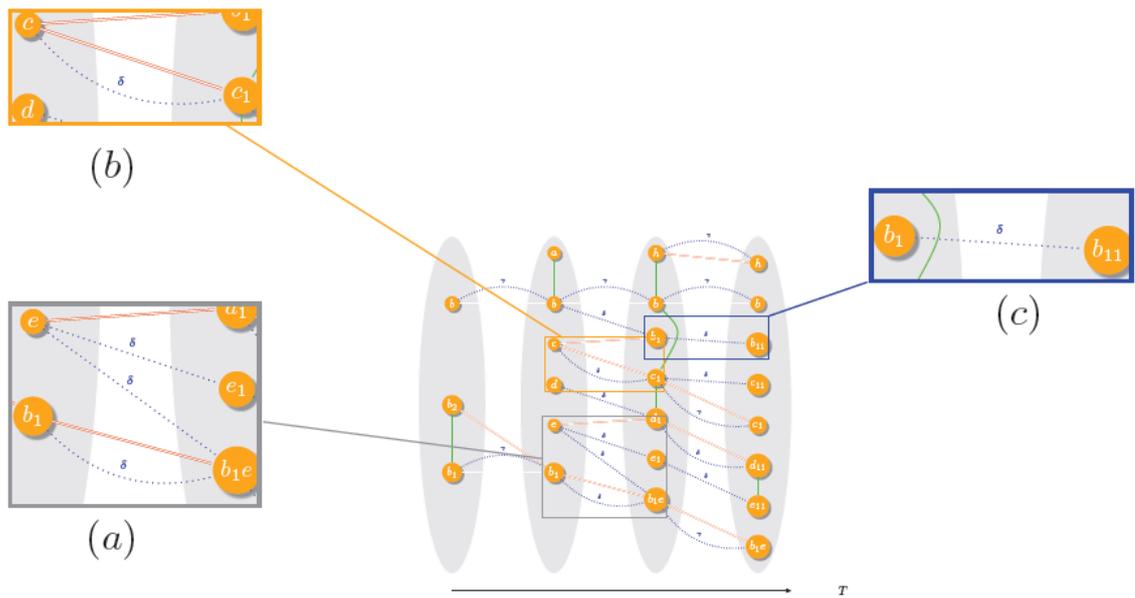


Figure 35 – Identification des processus de propagation dans l'évolution de la population de ronces : (a) La ronce e est greffée sur la ronce b_1 pour former b_1e ; (b) e_1 est issue soit d'un processus de marcottage ou bien de drageonnage de e ; (c) b_{11} est produite par dispersion des graines de la ronce b_1

Notre modèle spatio-temporel est basé sur les relations liées à l'identité des entités et à l'espace. Nous pouvons représenter ces relations à un instant ou bien dans le temps au travers des relations de filiations (identité), spatiales et spatio-temporelles. Cette modélisation générique sous forme de graphe nous permet d'envisager l'étude de phénomènes spatio-temporels très différents (e.g. déplacement de personnes, études de l'évolution de l'utilisation de parcelles de terrains). Il s'agit maintenant de définir les différentes fonctions (opérateurs) nous permettant d'accéder à l'information contenue dans le graphe et d'analyser le phénomène spatio-temporel sous-jacent.

II.2 Fonctions du modèle

On souhaite définir un certain nombre de fonctions nous permettant d'accéder à l'information contenue dans le graphe spatio-temporel, et d'étudier l'évolution des entités ainsi modélisée. Dans un premier temps, les fonctions sont définies sur les trois relations précédemment décrites (spatiale, spatio-temporelle et de filiation). Elles sont dédiées d'une part à la sélection des informations, et

d'autre part à l'identification des processus. Sur ce dernier point nous détaillons des processus classiques (e.g. naissance/mort d'une entité), et un exemple de processus lié à une application particulière (i.e. ici l'exemple de la section II.1). Ceci montre qu'il est a priori possible de spécifier tout opérateur nécessaire en fonction de l'application à traiter. Finalement, après avoir spécifié des opérateurs sur les entités d'un graphe spatio-temporel, on donne deux exemples de fonctions appliquées au graphe lui-même : la jointure de graphe et les routes. La jointure permet de mettre en relation des données provenant de différentes sources et ainsi d'enrichir les informations relatives au phénomène étudié. Les routes permettent de caractériser les chemins du graphe spatio-temporel.

II.2.1 Fonctions basées sur les relations

On définit le cadre formel général du type de fonction que l'on va utiliser.

II.2.1.1 Définition générale

Quelle que soit la relation choisie R sur un ensemble X d'éléments, on peut définir une fonction $R(x) \forall x \in X$:

$$R(x) = \{y \in X \mid xRy\} \quad (\text{II.1})$$

et $R : X \rightarrow \mathcal{P}(X)$.

On veut pouvoir composer la relation afin de tester les liens entre des sommets distants (dont la distance > 1) du graphe spatio-temporel.

$$R^d(x) = \{y \in X \mid xR^d y\} \quad (\text{II.2})$$

et $R^d : X \times \mathbb{N} \rightarrow \mathcal{P}(X)$, $d \geq 1$.

La relation $x R^d y$ représente la composition d -fold de R avec lui-même, c'est à dire par

exemple pour $d = 1$, $R^1 = R$ et pour $d > 1$, $R^d = R \bullet R^{d-1}$. L'ordre de la relation de composition⁷ \bullet est l'opposé de celui de la relation de composition traditionnelle \circ :

Soient deux relations Q et W définies sur l'ensemble X :

- $x(Q \circ W)y$ signifie qu'il existe un certain $z \in X$ tel que $x W z$ et $z Q y$
- $x(Q \bullet W)y$ signifie qu'il existe un certain $z \in X$ tel que $x Q z$ et $z W y$

L'ordre de la composition est significatif dès lors que la relation est transitive, et l'on souhaite que la relation générale R le soit. Figure 36, on remarque que pour posséder la propriété de transitivité il faut qu'il existe un z tel que $x Q z$ et $z W y$, et l'ordre de la composition doit suivre l'ordre (par exemple ici temporel) du graphe.

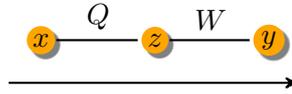


Figure 36 – On souhaite que les relations Q et W - et plus généralement R - soient transitives.

Nous allons appliquer ces deux fonctions génériques (II.1 et II.2) dans le cadre des relations que nous avons détaillées dans le chapitre précédent.

II.2.1.2 Fonctions basées sur la relation spatiale de connexion

Fonction de voisinage Il s'agit de la relation classique de voisinage en théorie des graphes. Les voisins directs d'un sommet sont ceux qui ont une arête adjacente à ce sommet. A un instant t , les relations spatiales de notre graphe spatio-temporel ne sont pas représentées orientées, mais une relation n'étant pas dans le cas général forcément symétrique, par souci de généricité, nous devons considérer l'orientation des relations. L'équation II.1 devient pour la fonction de voisinage spatial (Figure 37 (1)) :

$$\rho_s(x) = \{y \in \sigma \mid x \rho_s y\}$$

avec $\rho_s(x)$ l'ensemble des entités qui sont en relation spatiale ρ_s avec x .

Fonction de voisinage à distance d La relation spatiale n'étant pas toujours transitive, il est néanmoins intéressant de savoir si à défaut d'une relation spatiale explicite, il existe un lien entre

7. Dans la littérature ([Hirsch et Hodkinson, 2002]) utilise la notation ";", pour éviter toute confusion avec le symbole de ponctuation, nous lui avons préféré le symbole \bullet .

deux entités par l'intermédiaire d'une autre. La fonction de "voisinage à une certaine distance" le permet en déterminant les sommets qui sont à un certain nombre d'arêtes d'un sommet référence. Le calcul tient compte de l'orientation des arêtes (Figure 37 (2)). L'équation II.2 devient pour la fonction de voisinage spatial à une distance d :

$$\rho_s^d(x) = \{y \in \sigma \mid x \rho_s^d y\}$$

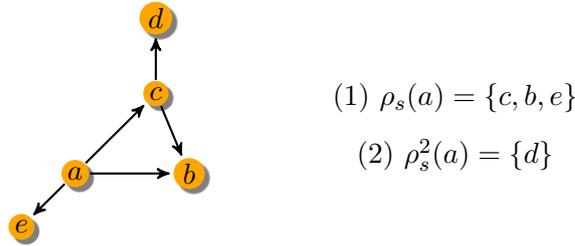


Figure 37 – (1) Fonction de voisinage spatial direct du sommet a ; (2) Fonction de voisinage spatial à une certaine distance, ici les voisins de a à la distance 2.

II.2.1.3 Fonctions basées sur la relation spatio-temporelle

Fonction de voisinage dans le temps Il s'agit d'une sémantique similaire à celle de la relation de voisinage à un instant t , mais l'ordre imposé par la dimension temporelle nous permet de spécifier certaines dénominations :

Une entité a est un **parent** spatio-temporel d'une autre entité b si $b \in \rho_{st}(a)$, et b est un **enfant** de a . Ceci signifie que a existe à un temps précédent de celui où se trouve b , et que leurs empreintes spatiales sont en relation spatiale ρ_s .

Afin de distinguer plus facilement la fonction relative aux parents et aux enfants, on re-définit l'équation II.1 qui devient pour les fonctions de voisinage spatio-temporel (Figure 38 (1)) deux fonctions :

$$\rho_{st}^-(x) = \{y \in \sigma \mid y \rho_{st} x\} \text{ si l'on cherche les parents de } x,$$

et

$$\rho_{st}^+(x) = \{y \in \sigma \mid x \rho_{st} y\} \text{ si l'on cherche les enfants de } x.$$

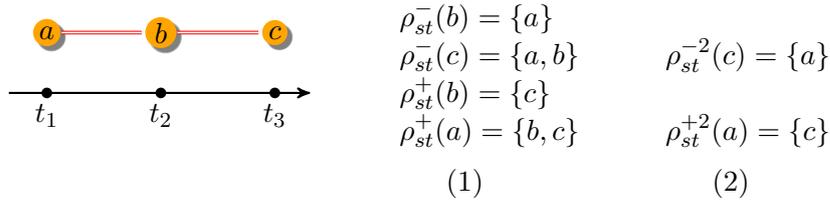


Figure 38 – (1) Fonctions de voisinage spatio-temporel ρ_{st}^- et ρ_{st}^+ ; (2) Fonctions de voisinage spatio-temporel ρ_{st}^{-d} et ρ_{st}^{+d} , $d \geq 1$.

Fonction de voisinage dans le temps à distance d Il s'agit de pouvoir déterminer les voisins spatio-temporels à une distance d fixée, aussi bien dans le passé et le futur :

Une entité a est un **ancêtre** spatio-temporel d'une autre entité b existant à un temps t , si $b \in \rho_{st}^d(a)$, $d \geq 1$, et b est appelé **descendant** de a . Ceci signifie que a existe au temps $t - d$ et que les empreintes spatiales de a et b sont en relation spatiale ρ_s

Afin de distinguer plus facilement la fonction relative aux ancêtres et aux descendants, on re-définit l'équation II.2 qui devient pour la fonction de voisinage spatio-temporel à une distance d (Figure 38 (2)) deux fonctions présentées ci-dessous. La dénomination d -step permet de faire explicitement référence à la distance qui sépare l'entité dont on veut connaître les ancêtres (respectivement les descendants), par rapport aux ancêtres (descendants) eux-mêmes :

$$\rho_{st}^{-d}(x) = \{y \in \sigma \mid y \rho_{st}^d x\} \text{ si l'on cherche les } d\text{-step ancêtres de } x,$$

et

$$\rho_{st}^{+d}(x) = \{y \in \sigma \mid x \rho_{st}^d y\} \text{ si l'on cherche les } d\text{-step descendants de } x.$$

II.2.1.4 Fonctions basées sur la relation de filiation

Nous reprenons les mêmes définitions que pour la relation spatio-temporelle en terme de dénomination (i.e. parent-enfant, ancêtre-descendant) et nous les transposons à la relation de filiation ρ_f . On ne fait pas de distinction entre les filiations de type continuation γ ou dérivation δ ⁸. Comme l'utilisation de ces fonctions est très similaire à celles relatives au spatio-temporel, on ne propose pas d'exemple spécifique, il suffit de remplacer la relation ρ_{st} par ρ_f dans la Figure 38.

Fonction de filiation dans le temps

8. Les définir revient à ne sélectionner que le type de filiation souhaité.

Une entité a est un **parent** en filiation d'une autre entité b si $b \in \rho_f(a)$, et b est un **enfant** en filiation de a . Ceci signifie que a existe à un temps précédent celui où se trouve b , et que b contient tout ou une partie de l'identité de a .

Afin de distinguer plus facilement la fonction relative aux parents et aux enfants, on re-définit l'équation II.1 qui devient pour les fonctions de filiation deux fonctions :

$$\begin{aligned} \rho_f^-(x) &= \{y \in \sigma | y \rho_f x\} \text{ si l'on cherche les parents de } x, \\ &\text{et} \\ \rho_f^+(x) &= \{y \in \sigma | x \rho_f y\} \text{ si l'on cherche les enfants de } x. \end{aligned}$$

Fonction de filiation dans le temps à distance d On veut pouvoir déterminer les voisins en filiation à une distance d fixée, aussi bien dans le passé et le futur :

Une entité a est un **ancêtre** en filiation d'une autre entité b existant à un temps t , si $b \in \rho_f^d(a)$, $d \geq 1$, et b est appelé **descendant** de a . Ceci signifie que a existe au temps $t - d$ et que b contient tout ou une partie de l'identité de a .

Afin de distinguer plus facilement la fonction relative aux ancêtres et aux descendants, on re-définit l'équation II.2 qui devient pour la fonction de voisinage de filiation à une distance d deux fonctions :

$$\begin{aligned} \rho_f^{-d}(x) &= \{y \in \sigma | y \rho_f^d x\} \text{ si l'on cherche les } d\text{-step ancêtres de } x, \\ &\text{et} \\ \rho_f^{+d}(x) &= \{y \in \sigma | x \rho_f^d y\} \text{ si l'on cherche les } d\text{-step descendants de } x. \end{aligned}$$

II.2.1.5 Application des fonctions basées sur les relations à l'exemple sur l'évolution de la population de ronces

Le graphe spatio-temporel de l'évolution de la population de ronces est rappelé Figure 39.

Nous nous plaçons dans le cadre d'une contamination. Nous étudions deux cas, le premier cherche à identifier le (ou les) processus de contamination, le second évalue les conséquences de ces processus.

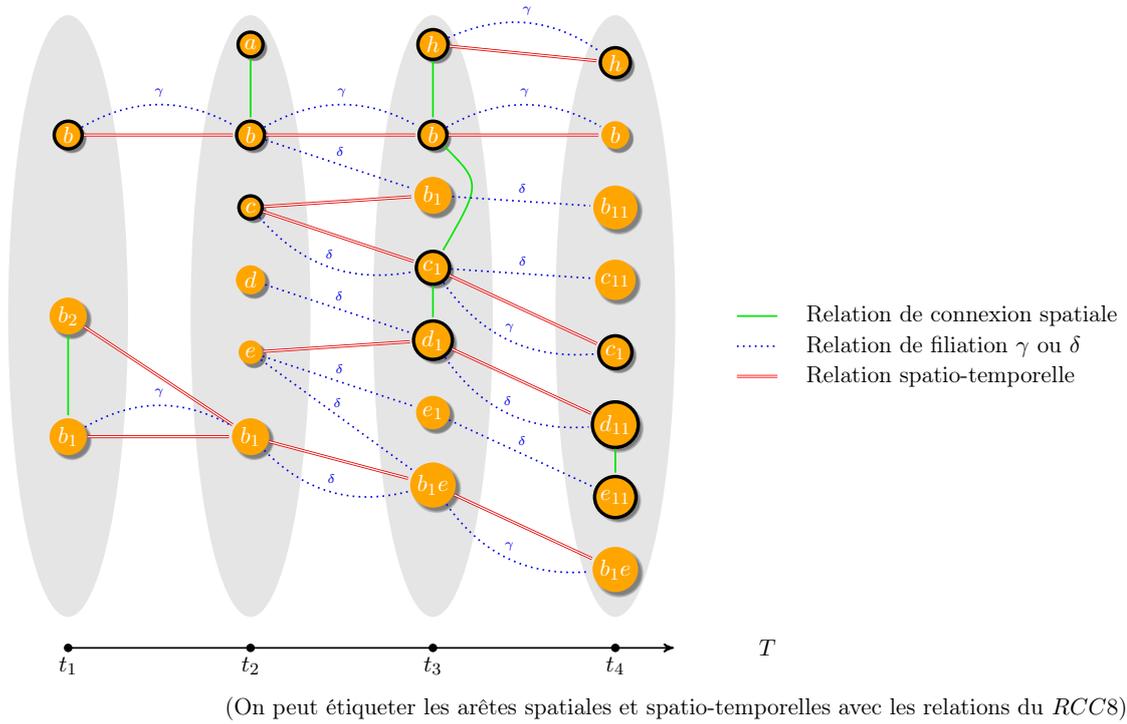


Figure 39 – Graphe spatio-temporel G_{ST} associé à l'évolution de la population de ronces. Les entités entourées en noir sont contaminées.

Cas 1 Il s'agit de trouver le processus de contamination à partir d'un ensemble d'entités contaminées connu. On suppose que les entités contaminées sont les suivantes :

Au temps t_1 : b

Au temps t_2 : a, b, c

Au temps t_3 : h, b, c_1, d_1

Au temps t_4 : h, c_1, d_{11}, e_{11}

On peut en déduire qu'une entité x est contaminée si et seulement si une entité y contaminée est telle que :

- $(y \in \rho_{st}^-(x) \wedge y \in \rho_f^-(x))$ avec ρ_f relation de dérivation.
ou bien
- $y \in \rho_s(x)$
ou bien
- $(y \in \rho_{st}^-(x) \wedge y \in \rho_f^-(x))$ avec ρ_f relation de continuation.

Ce qui correspond aux processus de drageonnage/marcottage ainsi qu'une simple relation spatiale pour les deux premiers. Ces deux processus sont ceux par lesquels survient la contamination. Le troisième cas suppose que la contamination perdure dans le temps et qu'une entité contaminée à un temps t restera contaminée au temps $t + 1$.

On peut remarquer que l'entité b_1 à t_3 n'est pas contaminée alors que c à t_2 l'est, et qu'il existe une relation spatio-temporelle entre les deux. Ceci est un exemple où distinguer relation spatiale et spatio-temporelle est utile. En effet, s'il est vrai que la contamination a lieu parce que deux ronces sont en relation spatiale, ce n'est pas pour autant parce qu'une ronce est en contact avec l'endroit où se trouvait une ronce contaminée qu'elle l'est à son tour.

Cas 2 Nous nous plaçons à nouveau dans le cas de contamination, mais nous considérons ici que les processus sont connus. Il s'agit de ceux identifiés dans le cas précédent, c'est-à-dire que la contamination a lieu lorsqu'une entité est créée par drageonnage/marcottage ou qu'il existe une relation spatiale entre deux entités. On sait également que la contamination perdure dans le temps sur une même entité.

Nous connaissons les entités contaminées au temps 4 : h, c_1, d_{11}, e_{11} .

L'algorithme suivant qui utilise les fonctions sur les relations nous permet d'identifier l'ensemble des entités qui ont été contaminées par les processus cités précédemment.

Algorithme : identification des entités contaminées

Soit E_i l'ensemble des entités contaminées au temps t_i

- (1) $i := 4$.
- (2) Trouver toutes les entités contaminées y telles que pour toute entité $x \in E_i$, $(y \in \rho_{st}^-(x) \wedge y \in \rho_f^-(x))$. On appelle cet ensemble d'entités contaminées $CSET$.
- (3) Pour chaque entité $y \in CSET$, trouver toutes les entités $z \in X(t_{i-1})$ telles que $z \in \rho_s^n(y)$, $\forall n \in \mathbb{N}$ tel que $\rho_s^n(y) \neq \{\}$. A la fin de cette étape, on ajoute à $CSET$ l'ensemble de ces entités z contaminées.
- (4) Si $i \neq 1$ alors $i := i - 1$, $E_i := CSET$ et retour au (2), sinon le résultat est $\cup_i(E_i)$.

On peut remarquer que la relation de continuation entre l'entité h à t_3 et h à t_4 (et pour toute entité en relation de continuation où l'ancêtre est contaminé) est prise en compte par l'algorithme

parce que les fonctions à notre disposition à ce stade ne différencient pas la dérivation de la continuation. Par conséquent l'algorithme considère qu'une continuation associée à une relation spatio-temporelle est équivalent aux processus drageonnage/marcottage (alors que normalement il devrait s'agir d'une relation spatio-temporelle associée à dérivation). Dans la pratique cela nous donne le bon résultat, car une entité contaminée reste contaminée, mais sémantiquement pour être tout à fait précis il faut bien être conscient que le test se fait à la fois sur la dérivation et sur la continuation.

Ces deux cas d'étude montrent qu'au-delà de l'aspect purement extraction de données, les fonctions relatives aux relations peuvent permettre d'identifier des processus dans un contexte particulier. Nous allons maintenant définir des fonctions permettant de retrouver directement les processus liés à l'application sur les ronces, et d'autres plus généraux dans le graphe.

II.2.2 Fonctions d'identification des processus

Processus classiques

Mort d'une entité : $\forall x \in \sigma$ et $\forall t \in T$, on définit la fonction $Mort : \sigma \times T \rightarrow \{Vrai, Faux\}$ qui renvoie Vrai si l'entité x , qui existe au temps t , est morte à $t + 1$:

$$Mort(x, t + 1) = \{Vrai \text{ si } x \notin \rho_f^1(x), Faux \text{ sinon}\}$$

Naissance d'une entité : $\forall x \in \sigma$ et $\forall t \in T$, on définit la fonction $Naissance : \sigma \times T \rightarrow \{Vrai, Faux\}$ qui renvoie Vrai si l'entité x , qui existe au temps t , est née à ce temps t :

$$Naissance(x, t) = \{Vrai \text{ si } x \notin \rho_f^-(x), Faux \text{ sinon}\}$$

Résurrection d'une entité : $\forall x \in \sigma$ et $\forall t \in T$, on définit la fonction $Resurrection : \sigma \times T \rightarrow \{Vrai, Faux\}$ qui renvoie Vrai si l'entité x , qui existe au temps t , n'existant plus au moins à $t - 1$, existe à nouveau à ce temps t :

$$Resurrection(x, t) = \{Vrai \text{ si } x \notin \rho_f^{-1}(x) \wedge x \in \rho_f^-(x), Faux \text{ sinon}\}$$

Processus lié à une application Dans l'exemple de la section II.1 on définit les processus de propagation de la ronce :

- (1) Création de nouvelle(s) ronce(s) fille(s) x_i par marcottage ou drageonnage : $x \delta x_i$ et $x \rho_{st} x_i$
- (2) Création de nouvelle(s) ronce(s) fille(s) x_i par dispersion des graines : $x \delta x_i$
- (3) Création de nouvelle(s) ronce(s) fille(s) xy par greffe de y sur x : $x \delta xy$ et $y \delta xy$

On peut aisément transformer ces propriétés en opérateurs, par exemple (3) devient :

- $\forall x, y$ deux ronces présentes à un instant $t \in T$, et xy présente à $t + 1$, on définit la fonction $Greffe : \sigma \times \sigma \times \sigma \times T \rightarrow \{Vrai, Faux\}$ qui renvoie vrai si xy est issue de la greffe de x et y :

$$Greffe(x, y, xy, t) = \{Vrai \text{ si } x \in \rho_f^{-1}(xy) \wedge y \in \rho_f^{-1}(xy) \wedge x \in \rho_{st}^{-1}(xy), \text{ Faux sinon}\}$$

- ou bien la fonction qui renvoie les parents responsables de la greffe de la ronce xy présente à $t \in T$:

$ParentsGreffe : \sigma \times T \rightarrow (\sigma \times \sigma)$, qui renvoie le couple des parents de xy :

$$ParentsGreffe(xy, t) = \{(x, y) | x \in \rho_f^{-1}(xy) \wedge x \in \rho_{st}^{-1}(xy) \text{ et } y \in \rho_f^{-1}(xy)\}$$

II.2.3 Fonction de jointure

Lors de l'étude d'un phénomène particulier, il est commun d'utiliser des sources de données complémentaires. Par exemple lors de l'étude du comportement d'une population, on peut disposer d'une base de données qui recense les personnes et leur liens de filiations, et d'une autre base qui comprend les lieux d'habitation de ces personnes et les relations spatiales qui existent entre ces lieux. Chacune des bases de données correspond à un graphe spatio-temporel. Il s'agit donc ici de combiner plusieurs de ces graphes pour obtenir un graphe résultat composé des informations provenant de toutes les sources de données initiales. La combinaison est effectuée grâce à la relation qui existe entre les entités présentes au sein de bases de données différentes. Dans l'exemple précédent, une relation possible est la relation spatiale qui existe dès lors qu'une personne habite dans un lieu.

Dans ce but, nous définissons trois fonctions de jointure : l'union, l'intersection et la jointure conditionnelle. Les deux premières sont relatives aux opérateurs ensemblistes classiques, la troisième effectue la combinaison de deux graphes selon un certain critère déterminé en fonction du contexte.

II.2.3.1 Fonctions d'union et d'intersection

Union : Soient deux graphes spatio-temporels $G_{ST_1} = (\sigma_1, E_1, T_1)$ et $G_{ST_2} = (\sigma_2, E_2, T_2)$. Leur union est égale à (Figure 40) :

$$G_{ST_1} \cup G_{ST_2} = (\sigma', E', T') \text{ avec } \sigma' = \sigma_1 \cup \sigma_2, E' = E_1 \cup E_2 \text{ et } T' = T_1 \cup T_2.$$

Intersection : Soient deux graphes spatio-temporels $G_{ST_1} = (\sigma_1, E_1, T_1)$ et $G_{ST_2} = (\sigma_2, E_2, T_2)$. Leur intersection est égale à (Figure 41) :

$$G_{ST_1} \cap G_{ST_2} = (\sigma', E', T') \text{ avec } \sigma' = \sigma_1 \cap \sigma_2, E' = E_1 \cap E_2 \text{ et } T' = T_1 \cap T_2.$$

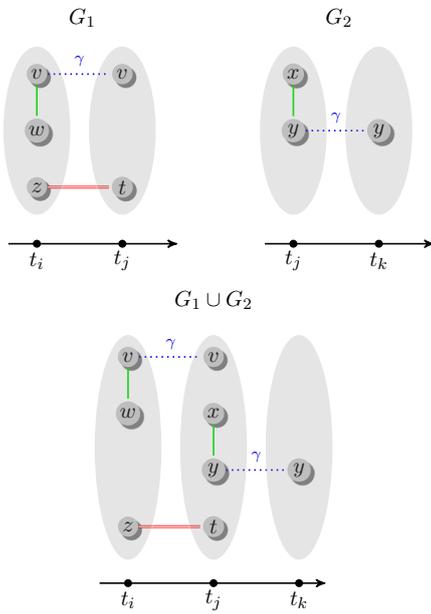


Figure 40 – Union G' de deux graphes spatio-temporels $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_1)$

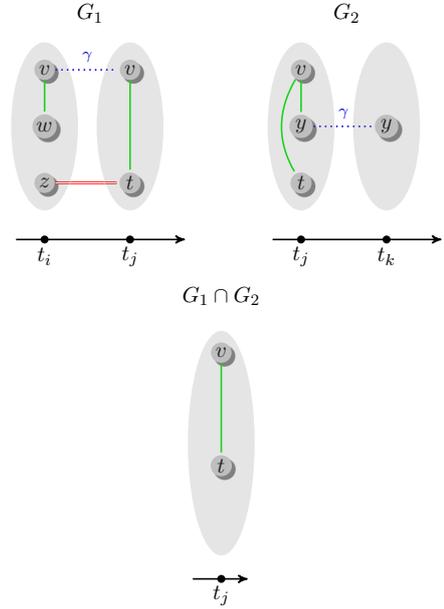


Figure 41 – Intersection G' de deux graphes $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_1)$

II.2.3.2 Fonction de jointure conditionnelle

Soient deux graphes spatio-temporels $G_{ST_1} = (\sigma_1, E_1, T_1)$ et $G_{ST_2} = (\sigma_2, E_2, T_2)$ de domaines de temps respectifs T_1 et T_2 . $X(t_i)$ et $X(t_j)$ sont les ensembles d'entités existantes aux temps $t_i \in T_1$ et $t_j \in T_2$, on pose $t_i \leq t_j$. La jointure conditionnelle entre ces deux graphes $G_{ST_1} \bowtie_{\rho} G_{ST_2}$ connecte les entités de G_{ST_1} et celles de G_{ST_2} au travers d'une relation $\rho \in \{\Gamma, \Phi\}$. La sémantique de la jointure conditionnelle dépend du type de la relation ρ (e.g. pour les bases de données personnes

et lieux d'habitation, ρ est relatif à une relation spatiale qui relie une personne à un lieu si cette personne y habite).

Remarque : Pour une relation ρ entre deux entités a et b , la notation $\rho[a, b]$ équivaut à écrire $a \rho b$. De plus, pour les jointures spatiale et spatio-temporelle on utilise une seule notation ρ_s dans la mesure où il s'agit du même ensemble de relations Γ . Il n'y a en effet aucun intérêt particulier à distinguer ces deux relations dans le cadre de la définition de la jointure conditionnelle.

La jointure spatiale : $G_{ST_1} \bowtie_s G_{ST_2} = (V', E', T')$ avec $V' = V_1 \cup V_2$, $T' = T_1 \cup T_2$ et $E' = E_1 \cup E_2 \cup \{(v, w) | t_i \in T' \wedge v \in X(t_i) \subseteq V_1 \wedge w \in X(t_i) \subseteq V_2 \wedge v \neq w \wedge \rho_s[v, w] \wedge \rho_s \in \Gamma\}$ (Figure 42).

La jointure spatio-temporelle : $G_{ST_1} \bowtie_{st} G_{ST_2} = (V', E', T')$ avec $V' = V_1 \cup V_2$, $T' = T_1 \cup T_2$ et $E' = E_1 \cup E_2 \cup \{(v, w) | t_i, t_j \in T' \wedge t_i < t_j \wedge v \in X(t_i) \subseteq V_1 \wedge w \in X(t_j) \subseteq V_2 \wedge \rho_s[v, w] \wedge \rho_s \in \Gamma\}$ (Figure 43).

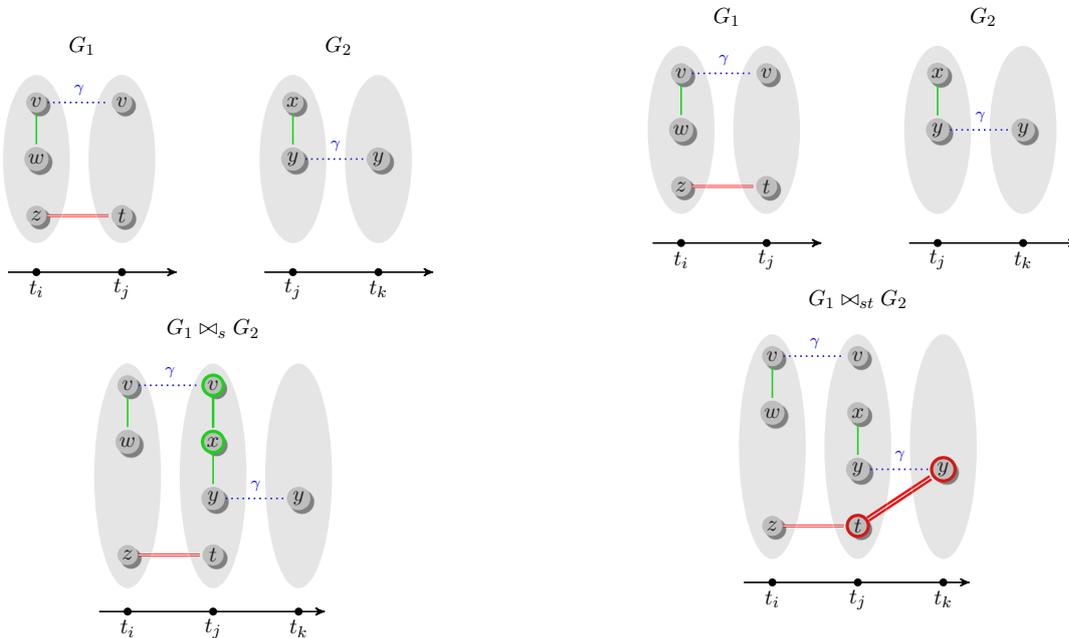


Figure 42 – Exemple de jointure spatiale $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_1)$ tel que $\rho_s[v, x]$, avec $v \in X(t_i) \subseteq V_1$, $x \in X(t_i) \subseteq V_2$.

Figure 43 – Exemple de jointure spatio-temporelle entre $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_2)$ tel que $\rho_s[t, y]$, avec $t \in X(t_i) \subseteq V_1$, $y \in X(t_j) \subseteq V_2$, $t_i < t_j$.

La jointure de filiation : $G_{ST_1} \bowtie_f G_{ST_2} = (V', E', T')$ avec $V' = V_1 \cup V_2$, $T' = T_1 \cup T_2$ et $E' = E_1 \cup E_2 \cup \{(v, w) | t_i \in T' \wedge v \in X(t_i) \subseteq V_1 \wedge w \in X(t_i) \subseteq V_2 \wedge v \neq w \wedge \rho_f[v, w] \wedge \rho_f = \delta$ (Figure 44).

La jointure de filiation temporelle : $G_{ST_1} \bowtie_{tf} G_{ST_2} = (V', E', T')$ avec $V' = V_1 \cup V_2$, $T' = T_1 \cup T_2$ et $E' = E_1 \cup E_2 \cup \{(v, w) | t_i, t_j \in T \wedge t_i < t_j \wedge v \in X(t_i) \subseteq V_1 \wedge w \in X(t_j) \subseteq V_2 \wedge \rho_f[v, w] \wedge \rho_f \in \Phi$.

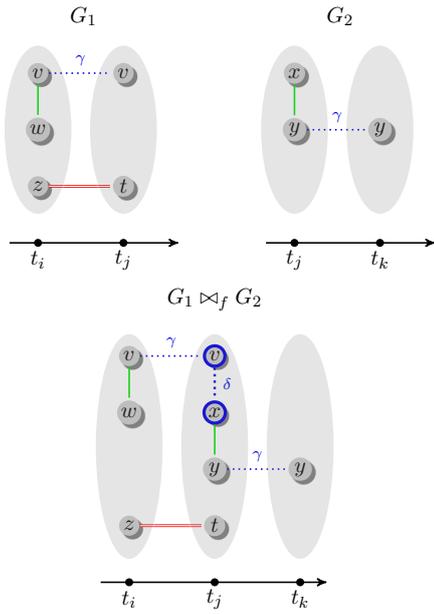


Figure 44 – Exemple de jointure de filiation entre $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_2)$ tel que $\rho_f[t, y]$, avec $t \in X(t_i) \subseteq V_1$, $y \in X(t_i) \subseteq V_2$, et ρ_f une relation de dérivation dans Φ .

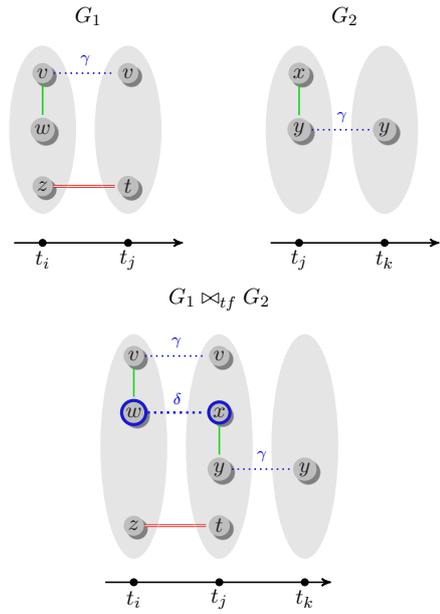


Figure 45 – Exemple de jointure de filiation temporelle, entre $G_1 = (V_1, E_1, T_1)$ et $G_2 = (V_2, E_2, T_2)$ tel que $\rho_f[w, x]$, avec $w \in X(t_i) \subseteq V_1$, $x \in X(t_j) \subseteq V_2$, $t_i < t_j$, et $\rho_f \in \Phi$.

Un exemple de jointure conditionnelle mettant en œuvre des domaines de temps se chevauchant est proposé Figure 46

II.2.3.3 Un exemple d'application de la jointure

Afin de donner un exemple comprenant des graphes spatio-temporels dont les entités sont de différents types, nous ne reprenons pas l'exemple sur les ronces précédemment introduit. Le phénomène à étudier ici est l'évolution des régions administratives au Chili entre 1925 et 2006. Cet autre exemple nous permet de souligner la généralité de notre modèle au vu des domaines d'applications très différents qu'il peut prendre en compte. On dispose de deux bases de données,

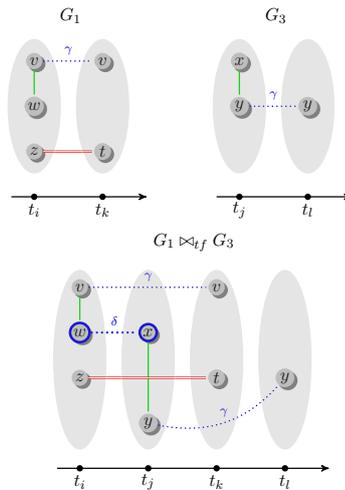


Figure 46 – Exemple de jointure de filiation temporelle pour des domaines de temps se chevauchant, i.e. $\exists t_i, t_k \in T_1$ et $\exists t_j, t_l \in T_2$ tels que $t_i < t_j < t_k < t_l$ et $\rho_f[w, x]$ avec $w \in X(t_i) \subseteq V_1$, $x \in X(t_j) \subseteq V_2$ et $\rho_f = \delta$.

l'une correspond à l'évolution des régions (Figure 47) et l'autre à l'évolution des provinces (48). Leur jointure spatiale (Figure 49) est définie quand une entité spatiale province est en relation spatiale avec une entité spatiale région. Elle a lieu aux temps où les deux graphes sont définis simultanément (i.e. 1976 et 2006). Le graphe résultant nous permet d'avoir une vision globale de la situation administrative, qui tient compte à la fois des régions et des provinces, et sur lequel il est possible d'appliquer toutes les fonctions que nous avons décrites précédemment.

Le graphe spatio-temporel modélise une évolution d'entités spatiales. La définition des différentes fonctions nous permet d'extraire l'information contenue dans ce graphe. La palette des fonctions que nous avons définies n'est pas exhaustive et il est possible d'en envisager beaucoup d'autres selon les besoins d'une application, en particulier sur les relations que nous avons définies ou sur d'autres (e.g. devant/derrière : cette entité est-elle derrière telle autre ?, liens sociaux : qui est l'ami de tel individu ?). Nous allons maintenant imaginer d'autres fonctions qui s'intéressent à la structure du graphe dans son ensemble, en s'appliquant cette fois au graphe tout entier et non à des entités particulières de celui-ci : les routes.

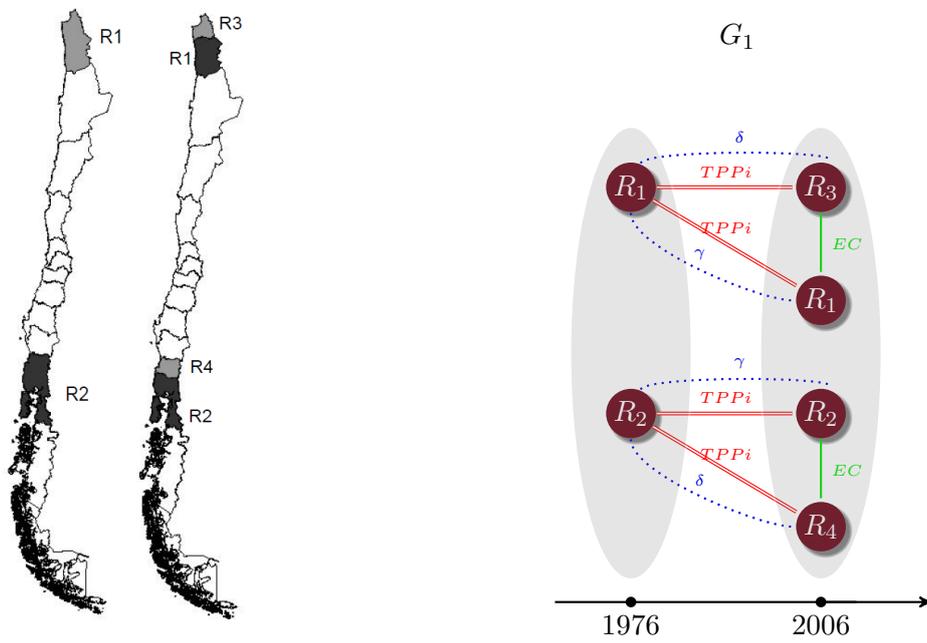


Figure 47 – Graphe spatio-temporel G_1 de l'évolution des régions. Les relations spatiales et spatio-temporelles sont étiquetées par les relations topologiques du $RCC8$.

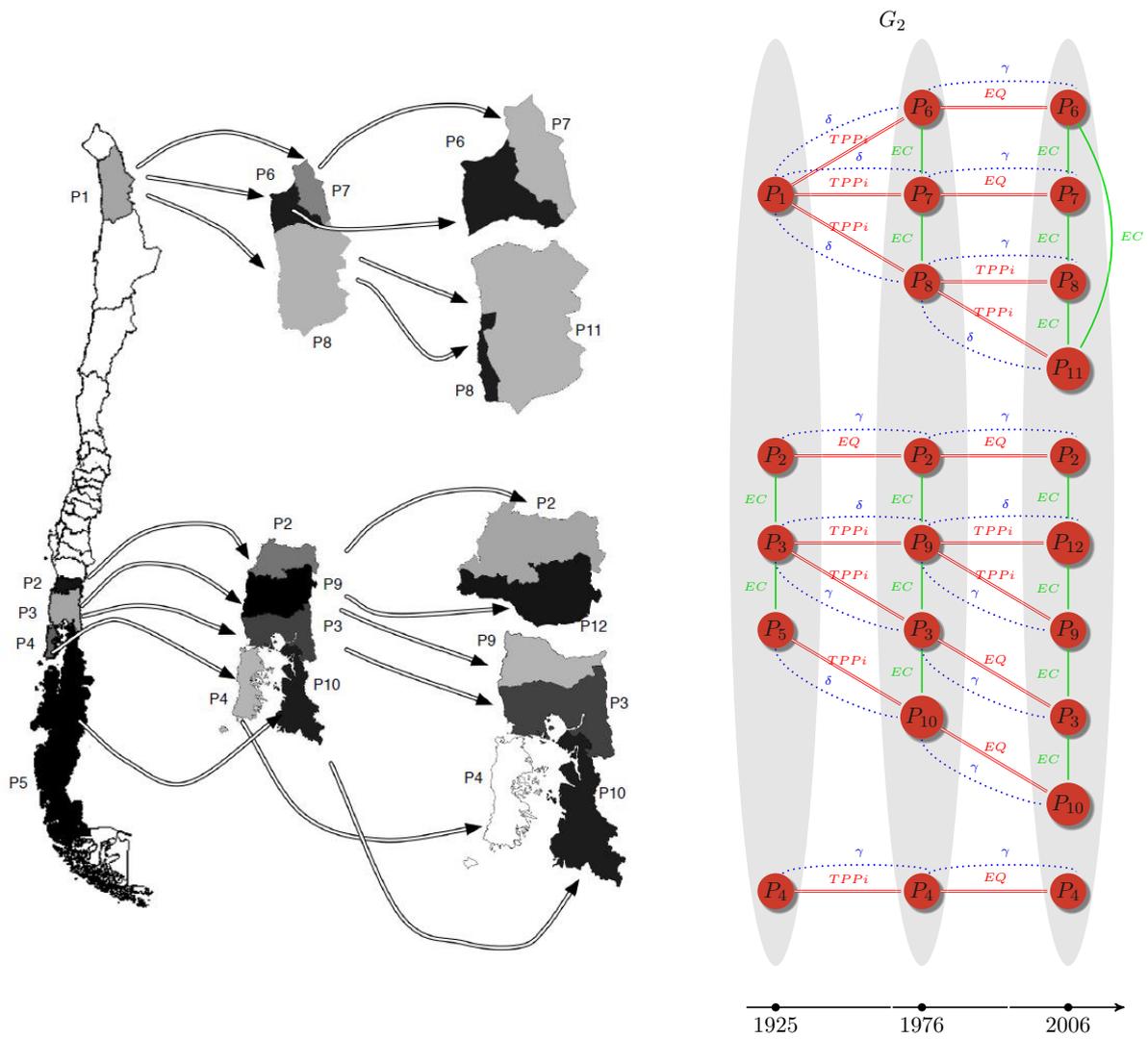


Figure 48 – Graphe spatio-temporel G_2 de l'évolution des provinces. Les relation spatiales et spatio-temporelles sont étiquetées par les relations topologiques du $RCC8$.

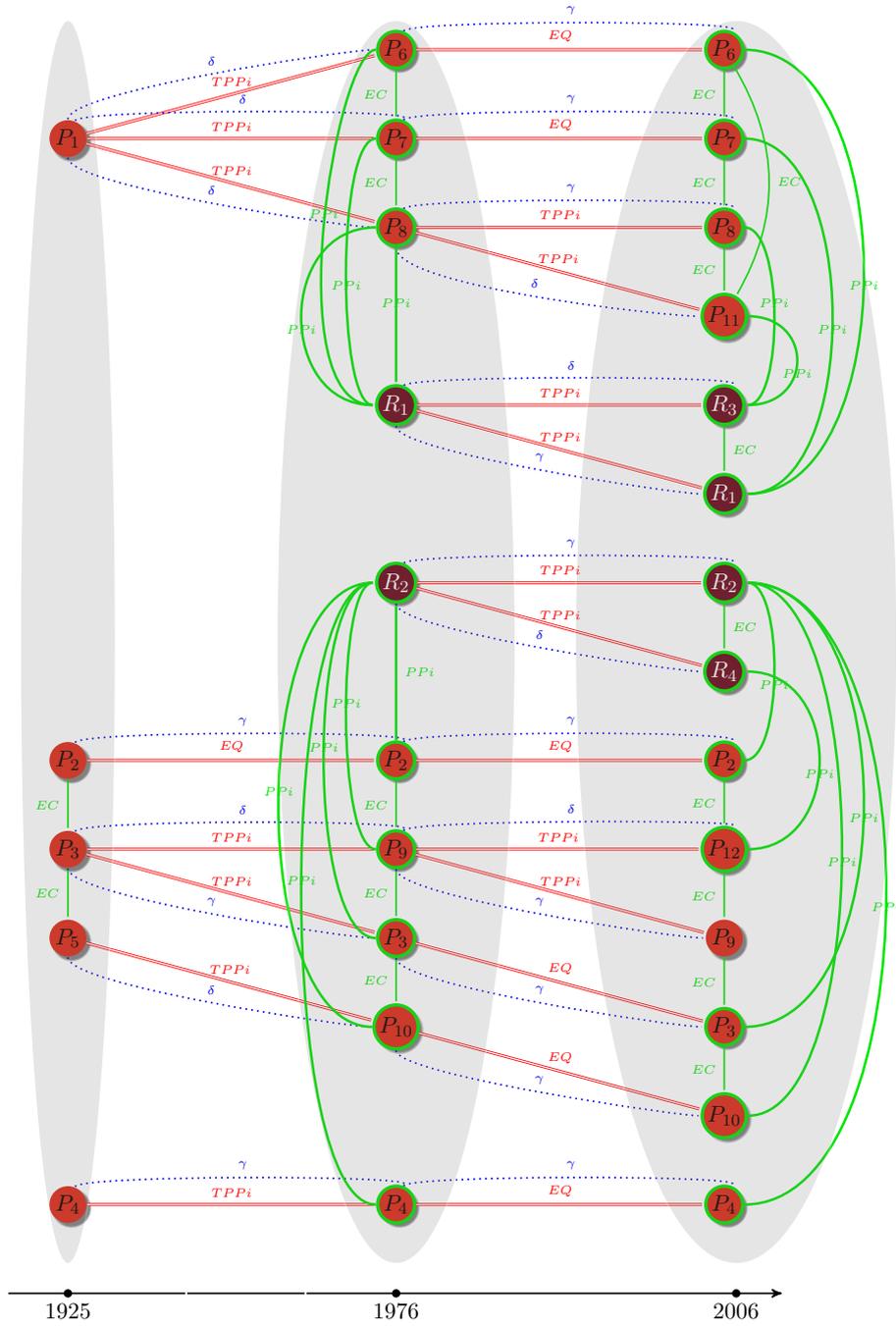


Figure 49 – Jointure spatiale entre provinces et régions au Chili.

II.2.4 Routes et relations

Au travers des fonctions décrites précédemment nous avons pu établir un certain nombre d'actions sur le graphe spatio-temporel, relatives aux relations spatiales, spatio-temporelles et de filiations, permettant d'extraire un certain nombre d'informations des données modélisées. Nous pouvons nous déplacer dans ce graphe (e.g. fonctions d'ascendance/descendance), sélectionner des entités selon certains critères (e.g. voisinage spatial), identifier des processus (e.g. Mort/Naissance d'entité), combiner différents graphes (e.g. opérateur de jointure), et toutes ces fonctions contribuent à l'analyse du phénomène modélisé. Dans cette dernière partie du chapitre, nous utilisons à nouveau les relations de base du modèle, mais cette fois nous souhaitons utiliser la composition de ces relations, par un nouvel outil : les routes. Nous nous attachons à qualifier la façon dont sont reliées les entités en suivant des chemins du graphe (i.e. suite d'entités reliées par des arêtes de relation), par exemple est-ce que le chemin de relations qui relie deux entités est unique ? S'il en existe plusieurs, proviennent-ils d'une seule source ? etc. Après la définition des différentes qualifications possibles de ces routes, nous donnons un exemple d'application basé sur la propagation des ronces.

II.2.4.1 Définition générale et routes spatio-temporelles

L'idée est que l'on peut suivre des chemins⁹ dans le graphe G_{ST} pour passer d'un temps t aux suivants et en retirer des informations sur l'enchaînement de leurs relations. Par composition de ces routes il va être possible de déduire des relations entre entités, entre des temps consécutifs ou non. On ne cherche pas ici à étudier les relations entre entités à un même temps.

Soit une relation R sur un ensemble d'éléments X . Pour un élément $A \in X$, on définit deux notions :

- Une route **qui part** d'une entité A est une séquence $A = A_0, A_1, \dots, A_n$ d'éléments de X tels que l'on ait $A_{i-1} R A_i, \forall i \in \{1, \dots, n\}$.
- Cette route **passe par** une entité $B \in X$ si $B = A_i$ pour un certain i .

Il en découle deux autres relations $\square R$ et $\diamond R$ sur X telles que, $\forall A, B \in X$:

$A \square R B$ si et seulement si **toutes** les routes **qui partent** de A **passent** par B .

$A \diamond R B$ si et seulement si il existe **une**¹⁰ route **qui part** de A et qui **passe** par B .

9. Un **chemin** dans un graphe est une suite (x_1, \dots, x_n) , $n \in \mathbb{N}$, de sommets telle que deux sommets consécutifs sont reliés par une arête.

10. Il existe d'autre(s) route(s) partant de A mais qui ne passent pas par B .

A partir de ces deux relations on définit $A \boxplus R B$ sur X qui permet d'assurer une relation où il existe au moins une route entre les deux. En effet, $A \square R B$ peut être vrai même si aucune route ne part de A . En d'autres termes, $A \boxplus R B$ signifie qu'il existe au moins une route partant de A et passant par B qui vérifie :

$$A \boxplus R B \text{ si et seulement si } (A \square R B) \wedge (A \diamond R B).$$

Nous complétons ces définitions grâce à la relation inverse R^{-1} , ainsi nous pouvons exprimer d'autres concepts :

$$B \square R^{-1} A \text{ si et seulement si toutes les routes qui arrivent à } B \text{ partent de } A.$$

$$B \diamond R^{-1} A \text{ si et seulement si il existe une route qui arrive à } B \text{ et qui part de } A.$$

Remarque : $\square R^{-1}$ équivaut à $\square(R^{-1})$ et non à $(\square R)^{-1}$. Ceci car $(\square R)^{-1}$ équivaut à dire que toutes les routes qui passent par B proviennent de A , seulement ce n'est pas exactement l'inverse de $\square R$, pour cela il faudrait s'assurer que ces routes sont les seules qui partent de A .

Plusieurs situations peuvent survenir entre A et B . En particulier, il est possible que (Représentation de ce cas Figure 50) :

- quelques routes partant de A passent par B ,
- mais que toutes les routes qui quittent A ne passent pas forcément par B ,
- et que toutes les routes qui passent par B proviennent de A

$$\text{Expression de ce cas : } (A \diamond R B) \wedge \neg(A \boxplus R B) \wedge (B \boxplus R^{-1} A).$$

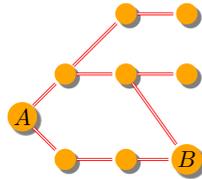


Figure 50 – Exemple de relation entre A et B conforme à la propriété $(A \diamond R B) \wedge \neg(A \boxplus R B) \wedge (B \boxplus R^{-1} A)$. Pour l'illustration on s'appuie sur R définie ici comme relation spatio-temporelle.

De manière générale il est possible d'exprimer la structure du lien qui existe entre deux entités A et B en utilisant les quatre propriétés suivantes :

- $A \diamond R B$
- $A \boxplus R B$
- $B \boxplus R^{-1} A$
- $B \diamond R^{-1} A$

Ces quatre propriétés ne sont pas indépendantes et il existe des liens d'implications entre elles,

que l'on peut synthétiser sur le diagramme Figure 51 où les flèches symbolisent l'implication. Par exemple, on voit bien que si toutes les routes qui partent de A passent par B ($A \boxplus R B$) alors c'est qu'il en existe au moins une ($A \diamond R B$).

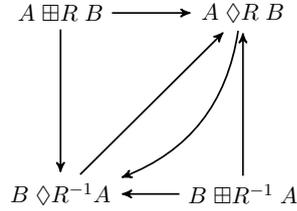
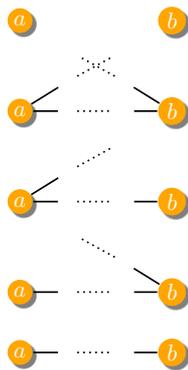


Figure 51 – Diagramme des relations d'implications entre les différentes propriétés utilisées pour qualifier les routes existant entre deux entités A et B .

Nous avons classé les cinq propriétés possibles entre A et B , et nous les avons regroupées selon une relation $\star R$. Cette relation peut prendre les valeurs AllAll, AllSome, SomeAll, SomeSome and None dans des expressions de type $A \star R B$, avec :

- AllAll $(A \boxplus R B) \wedge (B \boxplus R^{-1} A)$
- AllSome $(A \boxplus R B) \wedge \neg(B \boxplus R^{-1} A)$
- SomeAll $\neg(A \boxplus R B) \wedge (B \boxplus R^{-1} A)$
- SomeSome $(a \diamond R b) \wedge \neg(a \boxplus R b) \wedge \neg(b \boxplus R^{-1} a)$
- None $\neg(a \diamond R b)$

La signification des ces types de route est présentée Figure 52.



- None** : Il n'existe aucune route entre A et B .
- Some / Some** : Seules certaines des routes partant de A passent par B , et certaines des routes arrivant en B proviennent de A .
- Some / All** : Seules certaines des routes partant de A passent par B , mais toutes les routes arrivant en B proviennent de A .
- All / Some** : Toutes les routes partant de A arrivent en B , mais seules certaines passant par B proviennent de A .
- All / All** : Toutes les routes partant de A arrivent en B , et toutes les routes passant par B proviennent de A .

Figure 52 – Les différents types de routes

Si l'on considère la relation AllAll comme étant la plus restrictive et None comme celle qui ne caractérise pas le lien entre A et B (sinon par une absence de lien), on peut représenter ces cinq propriétés selon le treillis \mathcal{L} Figure 53.

Ce treillis nous permet de regarder s'il est possible de connaître l'ordre dans \mathcal{L} d'une relation l_3

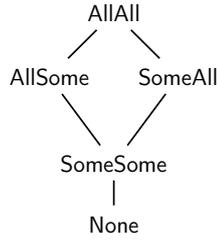


Figure 53 – Représentation des types de routes entre deux entités A et B sous forme de treillis

issue de la composition des relations l_1 et $l_2 \in \mathcal{L}$. C'est-à-dire tels que $(A l_1 B)$, $(B l_2 C)$ et $(A l_3 C)$. On peut affirmer que l'ordre le plus grand dans \mathcal{L} entre l_1 et l_2 restera inférieur ou égal à celui de l_3 .

Ceci peut être utile lorsque par exemple nous connaissons le type de relation $\star R$ entre A et B , et que nous le connaissons aussi entre B et C , nous avons alors des indications sur la relation entre A et C .

Remarque : Les routes pour identifier des processus On peut remarquer que si l'on se trouve entre deux temps consécutifs, certaines des relations de routes correspondent à des processus classiques de division (**SomeAll**) et de fusion (**AllSome**). Identifier un de ces deux types de route entre deux temps consécutifs revient donc à identifier un de ces deux processus. Par exemple, il est possible de définir une fonction $Split : \sigma \times \sigma \rightarrow \{Vrai, Faux\}$ qui renvoie *Vrai* si et seulement si $\forall x \in X(t_i), \forall y \in X(t_{i+1})$, on a $x \star R y = \text{SomeAll}$.

L'ensemble de ce que nous venons de décrire dans cette sous-partie II.2.4.1 pour la relation R fonctionne très bien pour R comme relation spatio-temporelle. La seule adaptation que nous avons à faire pour traiter les routes de filiations consiste à gérer les deux sous-relations γ et δ dans le treillis \mathcal{L} . On pose que l'ordre de la relation γ est plus grand que celui de δ dans \mathcal{L} et on définit l'ensemble des relations $\star \rho_f$ qui comprend sept relations analogues à celles de $\star R$ mais déclinées selon γ et δ . En effet, la transitivité de la relation de filiation nous permet de décrire les relations de composition qui existe entre deux entités se trouvant aux extrémités d'une route de filiation, et donc d'assigner l'étiquette de dérivation ou de continuation à la route selon le résultat de la composition.

- La composition de deux ou plusieurs relations de continuation renvoie une relation de continuation (et donc une route étiquetée γ).
- La composition d'une dérivation avec d'autres dérivations ou continuation renvoie une relation de dérivation (et donc une route étiquetée δ).

On établit ces règles de composition en considérant qu'à partir du moment où l'on a une

dérivation dans une route, l'identité de l'entité initiale par rapport à celles qui suivent la dérivation en est forcément affectée. Il s'agit donc globalement d'une dérivation.

D'où le treillis Figure 54, l'absence de la relation AllSome et SomeSome pour la continuation vient des contraintes que l'on a posées sur la relation de filiation.

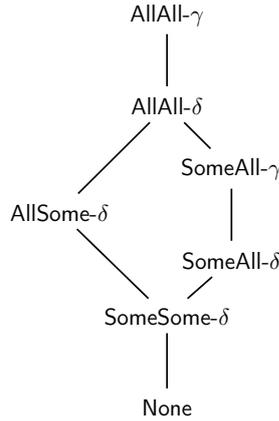


Figure 54 – Représentation des types de routes entre deux entités A et B sous forme de treillis pour la relation de filiation

II.2.4.2 Exemple d'utilisation des routes

On se base sur le graphe spatio-temporel relatif à la modélisation de l'évolution de la population de ronces, rappelé Figure 55.

On considère par exemple que la contamination s'effectue uniquement par le processus correspondant à la relation spatio-temporelle. Nous allons montrer comment l'utilisation de routes peut nous permettre d'identifier les entités contaminées, ou des propriétés relatives à la contamination.

Considérons deux ronces x et y respectivement aux temps t_i et t_j avec $i < j$:

- (a) On suppose que l'entité x est contaminée. Considérant un processus de contamination par relation spatio-temporelle, si $(x \star R y = \text{AllAll}) \vee (x \star R y = \text{SomeAll})$ alors toutes les entités sur ces routes sont contaminées. C'est par exemple le cas pour c et c_1 respectivement à t_2 et t_4 , ou b_1 et b_1e respectivement à t_1 et t_4 .

Réciproquement, il est possible d'inférer l'origine d'une contamination, si l'on a des routes de type AllAll ou SomeAll et si toutes les entités de ces routes sont contaminées. Par exemple, dès lors que l'on a $c \star R c_1 = \text{SomeAll}$, entre les temps t_2 et t_4 , on sait que la contamination provient de c . C'est aussi le cas entre les temps t_1 et t_4 pour $b_1 \star R b_1e = \text{AllAll}$, où b_1 serait

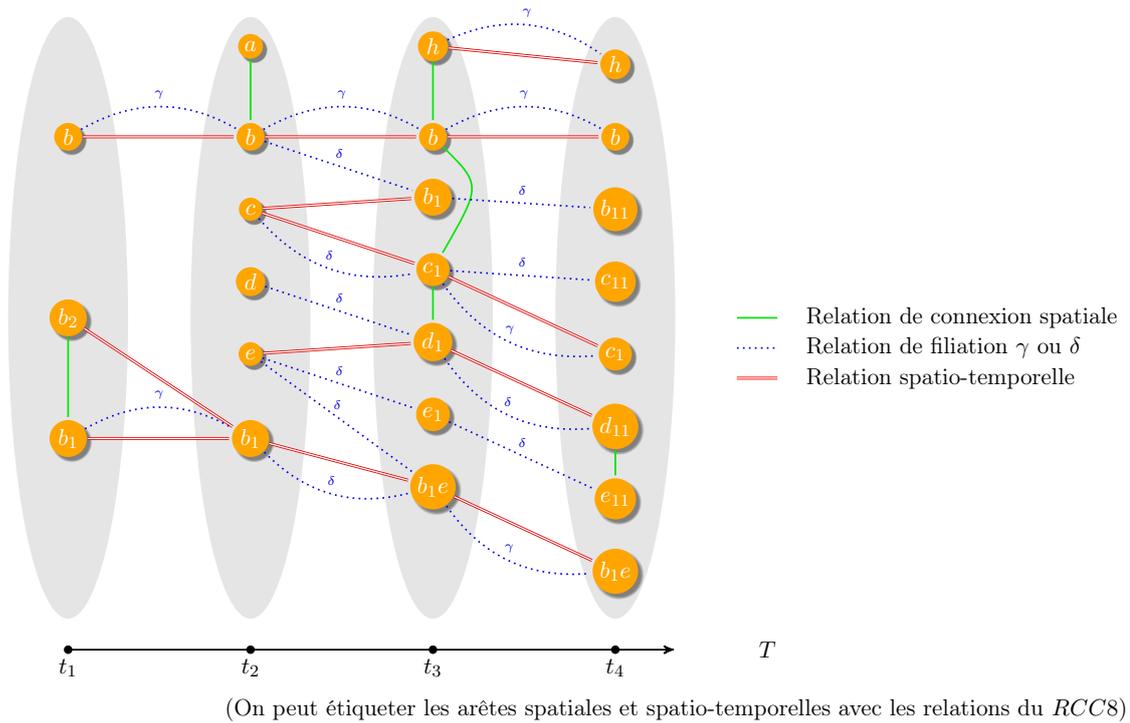


Figure 55 – Graphe spatio-temporel G_{ST} associé à l'évolution de la population de ronces.

à l'origine de la contamination.

- (b) Par contre, si l'on a $x \star R y = \text{AllSome}$ avec y que l'on sait contaminé, on ne peut pas être sûr que la contamination provienne de x . C'est par exemple le cas entre les temps t_1 et t_2 pour $b_2 \star R b_1 = \text{AllSome}$.
- (c) Pour illustrer la composition des filiations des routes concernées, on remarque que la route de filiation qui va de l'entité b au temps t_1 à l'entité b au temps t_4 n'est composée que de relation(s) de continuation. On peut donc étiqueter cette route par la relation de continuation γ , et ceci caractérise la relation de filiation qui existe entre ces deux versions de l'entité b . De la même façon, la route de filiation qui existe entre l'entité b à t_1 et l'entité b_{11} à t_4 est composée d'une relation de continuation et de deux dérivations. Cette route est donc étiquetée par une relation de dérivation δ , et ceci caractérise la relation de filiation qui existe entre b et b_{11} .

Nous avons montré par cet exemple la façon dont les fonctions relatives aux routes pouvaient être utilisées.

Ce chapitre décrit le modèle que nous avons mis en place pour l'analyse d'évolution d'entités dans l'espace et le temps. Nous avons défini formellement le graphe spatio-temporel au travers des trois relations (i.e. spatiale, spatio-temporelle, de filiation) que nous avons considérées comme étant nécessaires à l'analyse d'un phénomène spatio-temporel. Nous avons ensuite spécifié un certain nombre de fonctions qui permettent :

- d'extraire de l'information sur les entités du graphe par rapport aux trois relations (e.g. fonctions de voisinage),
- d'identifier des processus particuliers (e.g. fonctions de naissance/mort) ou liées à une application (e.g. greffe dans la propagation des ronces),
- de combiner plusieurs sources d'informations (e.g. jointure),
- de caractériser les relations entre entités par identification des différents types de chemins du graphe qui existent entre ces entités (i.e. route). Ces fonctions permettent aussi d'identifier des processus de division et de fusion entre deux temps consécutifs. Par ailleurs, étiqueter les routes de filiation par le résultat de la composition des relations de cette route, permet d'établir le type de la relation de filiation qui existe entre les entités présentes aux extrémités de cette route.

Dans le chapitre suivant nous allons nous attacher à spécifier un formalisme de contraintes qui nous permettra d'adapter ce modèle à des applications particulières, et de vérifier sa consistance au vu des contraintes définies.

Sommaire

III.1 Motivation	71
III.2 Description du formalisme pour la spécification des contraintes	73
III.2.1 Schéma général	74
III.2.2 Entité	75
III.2.3 Relations topologiques	77
III.2.4 Relations de filiations	77
III.2.5 Le graphe spatio-temporel	78
III.2.6 Fonctions de jointure	79
III.3 Spécification des contraintes	81
III.3.1 Les contraintes inhérentes au modèle : contraintes de domaine (<i>DC</i>)	81
III.3.2 Les contraintes dépendantes d'une application : contraintes sémantiques (<i>SC</i>)	84
III.4 Raisonner et vérifier la consistance	87
III.4.1 <i>Satisfiabilité</i> de l'ensemble de contraintes	87
III.4.2 Algorithme de consistance d'un ensemble de contraintes sur une instance de graphe	87

III.1 Motivation

Spécifier des contraintes sur un modèle équivaut à s'interroger sur la cohérence (i.e. *consistance*) de ses données, par rapport à la réalité qu'il est censé représenter. La formalisation de contraintes liées à la réalité que l'on souhaite modéliser, nous permet ainsi de tester la consistance des données relatives au modèle.

Plus généralement, la notion de consistance impose que les données modélisées satisfassent des

contraintes d'intégrité, imposées par les propriétés et les caractéristiques issues du modèle et/ou de l'application modélisée. En effet, [Rodríguez *et al.*, 2010] confirment qu'un état inconsistant est à proscrire dans l'ensemble des données qui instancie un modèle. Il survient dès lors que les données de la base violent un ensemble de contraintes d'intégrités (ICs). Cet état peut apparaître si un modèle doit gérer des données provenant de sources distinctes. De plus, même si ces sources sont localement consistantes, le résultat de leur intégration dans la base de données globale ne l'est pas forcément.

La représentation de telles contraintes d'intégrité concerne un nombre important d'applications à composante spatiale, où les contraintes représentées définissent la consistance des objets entre plusieurs vues consécutives ([Eckmann et Boulton, 2008, Sundarraj *et al.*, 2006]). La consistance spatio-temporelle peut se définir informellement comme une contrainte qui vérifie le caractère plausible de l'évolution d'une entité spatiale. Cette contrainte peut par exemple établir les conditions spatiales et sémantiques qui permettent de s'assurer que deux entités spatiales ont évolué (e.g. expansion, contraction) ou non. Au sein des SIG, ces contraintes sont utilisées afin de spécifier les relations entre entités spatiales représentées [Brisaboa *et al.*, 1998]. Ces contraintes peuvent se représenter à partir de règles composées d'un événement (la mise à jour des données d'une parcelle), d'une condition nécessaire à la réalisation de cet événement, et d'une action associée (la mise à jour des parcelles voisines) [Claramunt et Parent, 2003].

Des travaux récents ont abordé la consistance logique et la catégorisation de contraintes d'intégrité pour des entités spatiales et des relations spatio-temporelles floues [Shu *et al.*, 2006, Bejaoui *et al.*, 2007]. Ces approches sont utiles dans le contexte d'applications géographiques où les propriétés des entités observées et leurs évolutions ne sont pas forcément toujours précisément représentées ou connues. Elles définissent une contrainte spatio-temporelle (que l'on qualifiera ici de spatio-sémantique) comme une expression composée de conditions spatiales, temporelles et sémantiques. Ces conditions atomiques peuvent être combinées, produisant des conditions spatio-sémantiques, et des conditions spatio-sémantiques et temporelles. Un exemple de contrainte spatio-sémantique et temporelle donnerait : " deux espaces cultivables ne peuvent pas se chevaucher sur la même période de temps " [Bejaoui *et al.*, 2007].

A la différence des travaux précédents orientés vers la consistance d'entités spatio-temporelles et l'évolution de leurs propriétés, nous nous concentrons ici sur l'étude de la consistance de relations spatio-temporelles topologiques entre entités. C'est-à-dire la consistance relative à leur évolution, leurs transformations ou leurs modifications dues aux opérateurs de manipulation de données qui interagissent avec le modèle.

Par conséquent, dans le contexte de notre modèle de graphe spatio-temporel, lorsque les données de ce modèle viennent d'être initialisées ou mises à jour, il faut pouvoir s'assurer que le graphe spatio-temporel ainsi construit soit cohérent avec la réalité. Il est donc nécessaire de spécifier les contraintes qui sont relatives à cette réalité, et vérifier la consistance du modèle par rapport à ces contraintes.

Nous allons décrire deux types de contraintes d'intégrité **ICs** :

- (1) Notre modèle possède quelques contraintes que nous avons posées précédemment (Figure 28 page 41) et qui lui sont propres. Par exemple une relation de continuation entre deux entités ne peut avoir lieu à un même temps. Ces contraintes inhérentes au modèle sont appelées des contraintes de domaine (*Domain constraints DC*).
- (2) Dès lors que nous voulons adapter notre modèle à une application en particulier, il est possible d'établir des contraintes qui seront liées à cette application. Par exemple, dans le cadre de la propagation des ronces, comme il s'agit d'entités statiques, si une ronce existe à deux temps différents, il y a forcément une relation spatio-temporelle entre les deux. Il peut aussi s'agir de contraintes telles que "deux parcelles ne peuvent jamais se superposer", ou encore qu'"une entité de type bâtiment se trouve toujours dans une entité de type parcelle". Ces contraintes propres à une application sont appelées contraintes sémantiques (*semantic constraints SC*). Nous concentrerons les exemples sur des processus classiques comme la fusion ou la division d'entité(s), qui selon l'application doivent disposer d'une relation de filiation spécifique.

III.2 Description du formalisme pour la spécification des contraintes

Spécifier des contraintes sur un modèle impose de lever toute ambiguïté sur les constituants de ce modèle. Il nous faut donc préciser notre précédente formalisation pour disposer de tous les paramètres nécessaires lors de la description des contraintes. Nous appellerons la première formalisation proposée au chapitre II *modèle de graphe*, et celle que nous allons développer *modèle de données*. Les deux formalisations sont équivalentes et d'une certaine façon se complètent. Le modèle de graphe est de plus haut niveau et permet de raisonner en termes de problématique de graphe, le modèle de données,

plus technique, est à un niveau de détails plus fin¹. Par exemple, dans le modèle de graphe, nous avons posé comme hypothèse qu'une entité disposait d'une empreinte spatiale, étant entendu que nous connaissons les coordonnées de cette empreinte. Dans la description du modèle de données, nous nommons explicitement les paramètres géométriques d'une entité (i.e. $\mathcal{S} = \{S_1, \dots, S_n\}$ l'ensemble des attributs géométriques d'une entité qui prennent leurs valeurs dans $\mathcal{P}(\mathbb{R}^2)$).

Nous commençons par donner le schéma général de la formalisation du graphe spatio-temporel dans ce modèle de données, puis nous en détaillons les éléments principaux. Afin de montrer que cette formalisation est complètement équivalente à la première, nous terminons cette partie par un exemple de reformulation des fonctions de jointure.

III.2.1 Schéma général

Le schéma de données d'un graphe spatio-temporel est un tuple de la forme :

$$\Sigma = (\mathcal{U}, \mathcal{A}, \mathcal{T}, \mathcal{S}, \mathcal{I}, \mathcal{O}, \Phi, \mathcal{G}, \Gamma, \Theta, \Delta)$$

avec :

- (a) \mathcal{U} est le domaine (potentiellement infini) qui comprend les valeurs possibles des attributs.
- (b) \mathcal{A} est un ensemble d'attributs thématiques (non spatiaux) dont les valeurs sont prises dans $\mathcal{U} \cup \mathcal{I} \cup \mathcal{T}$.
- (c) $\mathcal{S} = \{S_1, \dots, S_n\}$ est un ensemble d'attributs géométriques d'une entité² qui prennent leurs valeurs dans $\mathcal{P}(\mathbb{R}^m)$. Pour ce qui nous concerne nous nous trouvons dans le cas où $m = 2$.
- (d) $\mathcal{T} = \{t_1, \dots, t_n\}$ est un ensemble fini partiellement ordonné d'instant à une granularité temporelle fixée.
- (e) \mathcal{I} est le domaine potentiellement infini de valeurs relatives à l'identité des entités. Elles identifient une entité à différents instants.
- (f) Φ est le domaine fini de valeurs qui identifient une relation de filiation (e.g. relation(s) de dérivation) .
- (g) \mathcal{O} est un ensemble fini de prédicats relatifs aux entités.

1. Un effet de bord à cette formalisation est qu'en donnant davantage d'information sur les différentes variables du modèle, elle prépare à une implémentation éventuelle du modèle et de ses contraintes.

2. Dans le cadre de ce chapitre, il faudrait parler davantage d'*objet* que d'*entité* puisque nous sommes en train de détailler la façon dont une entité est effectivement représentée dans la base de données, mais comme dès le départ nous avons choisi de ne pas différencier les deux termes, nous maintenons la dénomination d'*entité*.

- (h) \mathcal{G} est un ensemble de prédicats relatifs aux relations de filiations.
- (i) Γ est un ensemble de prédicats relatifs aux relations spatiales.
- (j) Θ est un ensemble d'opérateurs géométriques qui prennent des arguments spatiaux (e.g. *Aire, périmètre, union géométrique*).
- (k) Δ est un ensemble de prédicats qui s'appliquent aux valeurs thématiques d'attributs, comme des opérateurs de comparaisons (e.g. $<, >, =, \neq$). Seul l'opérateur de comparaison $' = '$ peut aussi s'appliquer dans le cas des opérateurs géométriques lorsque deux géométries d'entités sont égales.

Notations : La syntaxe du langage établie par le schéma Σ peut décrire le modèle de graphe, les fonctions que nous avons définies pour ce modèle, ainsi que les contraintes d'intégrité. Afin de simplifier la notation dans leurs spécifications, nous nous appuierons sur les notations ci-dessous :

- Les variables id, id_1, \dots dénotent les identifiants des entités spatio-temporelles, elles prennent leurs valeurs i, i_1, i_2, \dots , dans \mathcal{I} .
- t_1, t_2, \dots dénotent les attributs temporels, ils prennent leurs valeurs j, j_1, j_2, \dots dans \mathcal{T} .
- Les variables g, g_1, g', \dots dénotent les attributs géométriques dans \mathcal{S} , elles prennent leurs valeurs s_1, s_2, \dots dans $\mathcal{P}(\mathbb{R}^2)$.
- Les variables u, u', u_1, u_2, \dots dénotent les attributs thématiques dans \mathcal{A} , elles prennent leurs valeurs c, c', c_1, c_2, \dots dans \mathcal{U} .
- Les symboles $\bar{u}, \bar{u}', \bar{u}_1, \bar{u}_2, \dots$ et les symboles $\bar{c}, \bar{c}_1, \bar{c}_2, \dots$, dénotent une séquence potentiellement vide de variables et valeurs distinctes respectivement.
- Les valeurs δ, δ', \dots sont prises dans Φ .
- Le symbole $\bar{\forall}$ nous permet -quand il n'y a pas d'ambiguïté possible- de donner une équation en considérant que tous les paramètres ont bien été spécifiés (e.g. $\bar{\forall} g(x)$ au lieu de $\forall x \in \mathbb{R} g(x)$).

III.2.2 Entité

Une entité est représentée par un prédicat spatio-temporel $O \in \mathcal{O}$. Il est de la forme :

$$\boxed{O(id, t, u_1, \dots, u_n, g)}$$

avec :

- id est un attribut thématique (identité) dans \mathcal{A} qui permet d'identifier une entité, il prend

ses valeurs dans \mathcal{T} .

- t est un attribut thématique (attribut temporel) dans \mathcal{A} dont les valeurs sont des instants temporels pris dans \mathcal{T} .
- $\langle u_1, \dots, u_n \rangle \in \mathcal{A}^n$ est une séquence de valeurs possibles pour les attributs thématiques³.
- g est un attribut spatial ou géométrique⁴ dans \mathcal{S} .

Un prédicat représente l'entité à tous les instants où elle existe. Il faut être vigilant sur le fait qu'un prédicat O n'est pas équivalent à une entité à un instant t , c'est-à-dire qu'un prédicat O_i ne représente pas forcément l'entité d'identifiant id_i . Par exemple on peut définir deux prédicats $O_1(id_1, t_1, \dots)$ et $O_2(id_2, t_1, \dots)$: il s'agit de deux entités qui existent au temps t_1 . Si l'entité d'identifiant id_1 perdure en t_2 , on aura le prédicat $O_1(id_1, t_2, \dots)$.

Une instance du prédicat $O(id, t, u_1, \dots, u_n, g)$ est un ensemble fini de tuples de la forme $O(i, j, c_1, \dots, c_n, s)$, où $\langle i, j \rangle$ identifie l'entité d'identité id au temps j ; c_1, \dots, c_n sont les valeurs des attributs thématiques possibles dans \mathcal{U}^n , et s est la géométrie de l'entité dans \mathcal{S} (s prend ses valeurs dans $\mathcal{P}(\mathbb{R}^2)$).

Dans ce schéma, nous ne faisons pas explicitement de distinction entre les différents types d'entités (e.g. parcelle, personne), ceci peut être géré en passant par un attribut thématique qui aurait pour valeur le type de l'entité.

Par contre, nous distinguons explicitement l'identité, elle permet d'identifier de façon unique une entité à n'importe quel instant temporel. Comme une entité peut exister à plusieurs instants différents, les attributs $\langle id, t \rangle$ identifient complètement une entité d'identifiant id et d'attribut temporel t .

3. Cette séquence peut-être vide

4. Ce modèle ne considère qu'une seule géométrie possible par entité, mais on peut imaginer étendre cette définition à plusieurs.

III.2.3 Relations topologiques

On considère que chaque prédicat $\Upsilon \in \Gamma$ représente une relation topologique qui possède sa propre sémantique (e.g. l'entité A touche l'entité B : il existe au moins un point qui est commun à l'empreinte spatiale de A et à celle de B .), et qu'il est de la forme :

$$\Upsilon(s_1, s_2)$$

où s_1 et s_2 sont des valeurs géométriques prises dans $\mathcal{P}(\mathbb{R}^2)$. Comme décrit dans le chapitre II, nous nous concentrons sur les relations topologiques du *RCC8*, mais il est bien entendu que n'importe quel ensemble de relations topologiques binaires peut être utilisé. On peut remarquer par ailleurs que si les relations topologiques sont explicitement représentées dans le graphe spatio-temporel par des arêtes, dans le schéma sous-jacent proposé ici, elles sont calculées grâce à la géométrie des entités. Ainsi, du point de vue de ce schéma les relations spatiales et spatio-temporelles sont équivalentes, ce qui est assez cohérent puisqu'il s'agit de la même relation. Pouvoir, du point de vue du graphe, identifier de manière explicite l'une ou l'autre est nécessaire, mais on convient que d'un point de vue technique il est normal que les deux se confondent. Ce choix simplifie grandement la preuve de consistance de l'ensemble de contraintes sémantiques qui a été effectuée dans l'article [Del Mondo *et al.*, 2011].

III.2.4 Relations de filiations

Chaque prédicat de filiation $R \in \mathcal{G}$ est de la forme :

$$R(id_1, t_1, \rho, id_2, t_2)$$

où les attributs id_1 et id_2 sont relatifs à l'identité, t_1 et t_2 sont des attributs temporels, et ρ est un attribut de filiation (i.e. type de la filiation) qui prend ses valeurs dans Φ .

Une instance de ce prédicat de filiation $R(id_1, t_1, \rho, id_2, t_2)$ est un ensemble fini de tuples $R(i_1, j_1, \rho_f, i_2, j_2)$ où $\langle i_1, j_1 \rangle$ et $\langle i_2, j_2 \rangle$ prennent leurs valeurs dans $\mathcal{I} \times \mathcal{T}$, et ρ_f est une relation de filiation dans Φ .

La relation de continuation γ peut être dérivée des prédicats relatifs aux entités si leurs attributs relatifs à l'identité ($\in \mathcal{I}$) sont identiques. La relation de dérivation δ nécessite par contre d'être

spécialisée⁵ en fonction de l'application modélisée. En effet, il existe des relations de dérivation relatives à différents processus (e.g. fusion, séparation) que nous introduirons lors de la spécification des contraintes relatives à ces mêmes processus (Section III.3.2).

L'existence du prédicat R implique l'existence des prédicats associés aux entités concernées par R , correctement instanciés. Cette contrainte d'intégrité sera formalisée comme une contrainte de domaine (Equation III.2).

III.2.5 Le graphe spatio-temporel

Un graphe spatio-temporel G_{ST} dans le modèle de données Σ est un couple (V, E) avec :

- $V = \{O_1, \dots, O_n\}$ l'ensemble des prédicats relatifs aux entités tel que $O_i \in \mathcal{O}, \forall i \in [n]$.
- $E = \{R_1, \dots, R_m\}$ l'ensemble des prédicats de filiation tels que pour tous $R_i \in \mathcal{G}, \forall i \in [m]$, de la forme $R_i(id_1, t_1, \rho, id_2, t_2)$, id_1, t_1 et id_2, t_2 sont des attributs dans O_i et O_j , respectivement, qui sont tous deux dans V .

Une instance de ce graphe spatio-temporel dans ce modèle de données est un ensemble fini d'instances $O_i \in V$ et $R_i \in E$. La correspondance du modèle de graphe avec ce schéma revient à transformer :

- les sommets du graphe en tuples de prédicats relatifs aux entités,
- les arêtes de filiation du graphe en tuples de prédicats relatifs aux filiations,

Un exemple de correspondance est proposé Figure 56, il reprend les données relatives aux régions du Chili introduites lors de la définition des jointures dans le modèle de graphe (Section II.2.3.3). Dans le schéma de données Σ de cette figure, le domaine de temps \mathcal{T} est un ensemble d'années $\{1976, 2006\}$ et l'ensemble des relations de filiations prend ses valeurs dans $\Phi = \{\delta\}$. Le graphe spatio-temporel est composé de prédicats relatifs aux entités $Region(id, year, name, geometry)$, où $\langle id, year \rangle$ identifient de manière unique la région d'identifiant id et d'attribut temporel $year$. Un attribut thématique $name$ renseigne le nom de la région, et l'attribut à valeur géométrique $geometry$ nous renseigne sur la géométrie de cette région. Le schéma contient également les relations de filiations de type dérivation sous la forme du prédicat $R_{RR}(id_1, year_1, \rho, id_2, year_2)$.

Les relations spatiales du graphe ne sont pas stockées dans ce modèle de données, mais dérivées des attributs géométriques des entités considérées⁶. De même, les relations de filiation du type

5. A moins que l'on ne connaisse pas la sémantique exacte de la relation de filiation pour une certaine application.

6. Le passage du modèle de données vers le modèle de graphe nécessite donc de re-calculer toutes les relations topologiques pour déterminer l'ensemble des arêtes spatiales et spatio-temporelles du graphe. Cet aspect est discuté

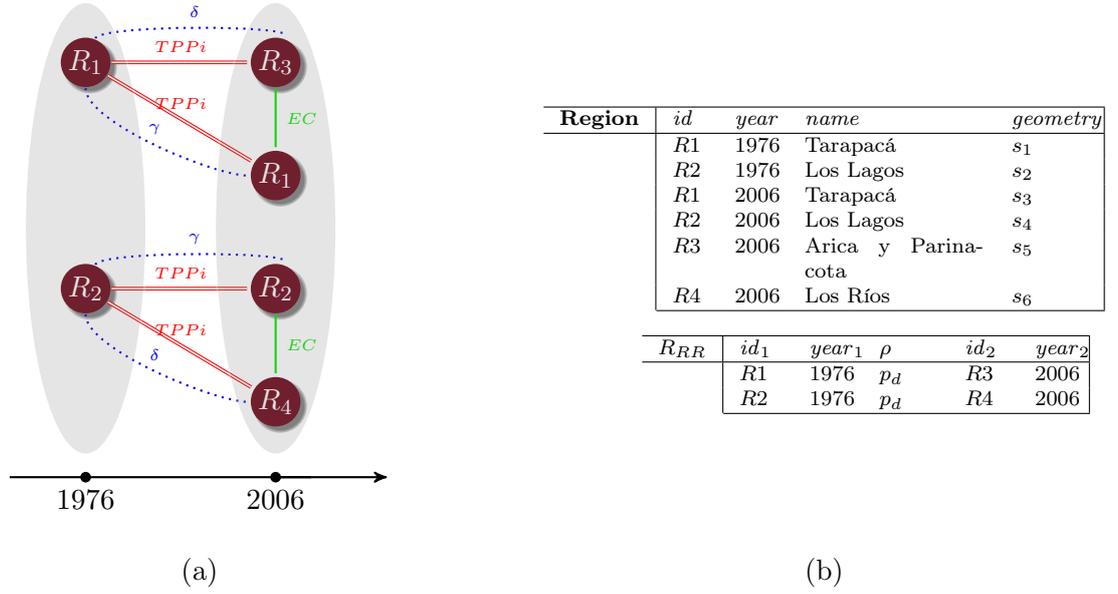


Figure 56 – Correspondance entre (a) un graphe spatio-temporel et (b) son instantiation comme schéma de données Σ .

continuation sont dérivées des attributs relatifs à l'identité des entités considérées (i.e. valeurs d'attributs identiques).

III.2.6 Fonctions de jointure

Afin de montrer que la formalisation dans le modèle de données peut représenter l'ensemble de ce qui a été défini dans le modèle de graphe, nous proposons ici un exemple avec les fonctions de jointure que nous formalisons pour le modèle de données.

Soient D_1 et D_2 des instances des graphes spatio-temporels $G_{ST1} = (V_1, E_1)$ et $G_{ST2} = (V_2, E_2)$, respectivement.

Union : L'union de deux graphes D_1 et D_2 correspond à un graphe dont l'ensemble des entités $\{O(i, j, \bar{c}, s)\}$ et l'ensemble des arêtes $\{R(i_1, j_1, \delta, i_2, j_2)\}$ sont définis tels que :

$$D_1 \cup D_2 = (\{O(i, j, \bar{c}, s) \mid O(i, j, \bar{c}, s) \in V_1 \vee O(i, j, \bar{c}, s) \in V_2\}, \\ \{R(i_1, j_1, \delta, i_2, j_2) \mid R(i_1, j_1, \delta, i_2, j_2) \in E_1 \vee R(i_1, j_1, \delta, i_2, j_2) \in E_2\})$$

dans l'annexe relative à l'implémentation du modèle.

Intersection : L'intersection de deux graphes D_1 et D_2 correspond à un graphe dont l'ensemble des entités $\{O(i, j, \bar{c}, s)\}$ et l'ensemble des arêtes $\{R(i_1, j_1, \delta, i_2, j_2)\}$ sont définis tels que :

$$D_1 \cap D_2 = (\{O(i, j, \bar{c}, s) | O(i, j, \bar{c}, s) \in V_1 \wedge O(i, j, \bar{c}, s) \in V_2\}, \\ \{R(i_1, j_1, \delta, i_2, j_2) | R(i_1, j_1, \delta, i_2, j_2) \in E_1 \wedge R(i_1, j_1, \delta, i_2, j_2) \in E_2\})$$

La jointure de filiation ou filiation temporelle : La jointure de filiation, temporelle ou non, (tout dépend des attributs temporels que l'on choisit) de deux graphes D_1 et D_2 correspond à un graphe dont l'ensemble des entités $\{O(i, j, \bar{c}, s)\}$ et l'ensemble des arêtes $\{R(i_1, j_1, \delta, i_2, j_2)\}$, ainsi que l'ensemble des relations de filiation $\{R(i_1, j_1, \delta, i_2, j_2)\}$ (relations ajoutées par la condition de jointure) sont définis tels que :

$$D_1 \bowtie_f D_2 = (\{O(i, j, \bar{c}, s) | O(i, j, \bar{c}, s) \in V_1 \vee O(i, j, \bar{c}, s) \in V_2\}, \\ \{R(i_1, j_1, \rho_d, i_2, j_2) | R(i_1, j_1, \rho_d, i_2, j_2) \in E_1 \vee R(i_1, j_1, \rho_d, i_2, j_2) \in E_2\} \cup \\ \{R(i_1, j_1, \rho_d, i_2, j_2) | \exists i_1 i_2, j_1 j_2, \bar{c}_1 \bar{c}_2, s_1 s_2 (O(i_1, j_1, \bar{c}_1, s_1) \in V_1 \wedge O(i_2, j_2, \bar{c}_2, s_2) \in \\ V_2 \wedge \exists \delta \text{ entre } \langle i_1, j_1 \rangle \text{ et } \langle i_2, j_2 \rangle)\})$$

La jointure spatiale ou spatio-temporelle : Comme les relations spatiales ne sont pas explicitement gérées dans le modèle de données, la jointure spatiale (ou spatio-temporelle) équivaut à une union de graphe :

$$D_1 \bowtie_s D_2 = D_1 \cup D_2$$

Ceci signifie que dès lors que l'on fait une jointure sur la relation spatiale et que l'on part du modèle de données vers le modèle de graphe, il faut calculer toutes les relations spatiales potentielles entre les entités de chacun des deux graphes. Le graphe résultat de la jointure contiendra les relations spatiales (ou spatio-temporelles) qui existent effectivement, mais celles-ci n'apparaissent pas explicitement au départ dans le modèle de données résultat de la jointure.

Un exemple de jointure spatiale est proposé Figure 57a. On considère D_1 l'instance de graphe relative aux régions et D_2 celle relative aux provinces. On les retrouve respectivement dans les Figures 47 (dont le modèle de données a été précisé Figure 56 et Figure 48. Le schéma de données de l'évolution des provinces est précisé Figure 57b : il contient les prédicats relatifs aux entités $Province(id, year, name, geometry)$ et des prédicats de filiation $R_{PP}(id_1, year_1, \rho, id_2, year_2)$ avec ρ relation de filiation entre provinces. Le domaine de temps \mathcal{T} du graphe résultat de la jointure

ture correspond à l'union des domaines de temps associés aux graphes D_1 ($\{1976, 2006\}$) et D_2 ($\{1925, 1976, 2006\}$), et on a $\Phi = \{\delta\}$.

Concrètement l'instanciation de la jointure spatiale dans le schéma de données Σ est composée des prédicats relatifs aux entités *Region* et *Province*, ainsi que des prédicats de filiation R_{RR} et R_{PP} .

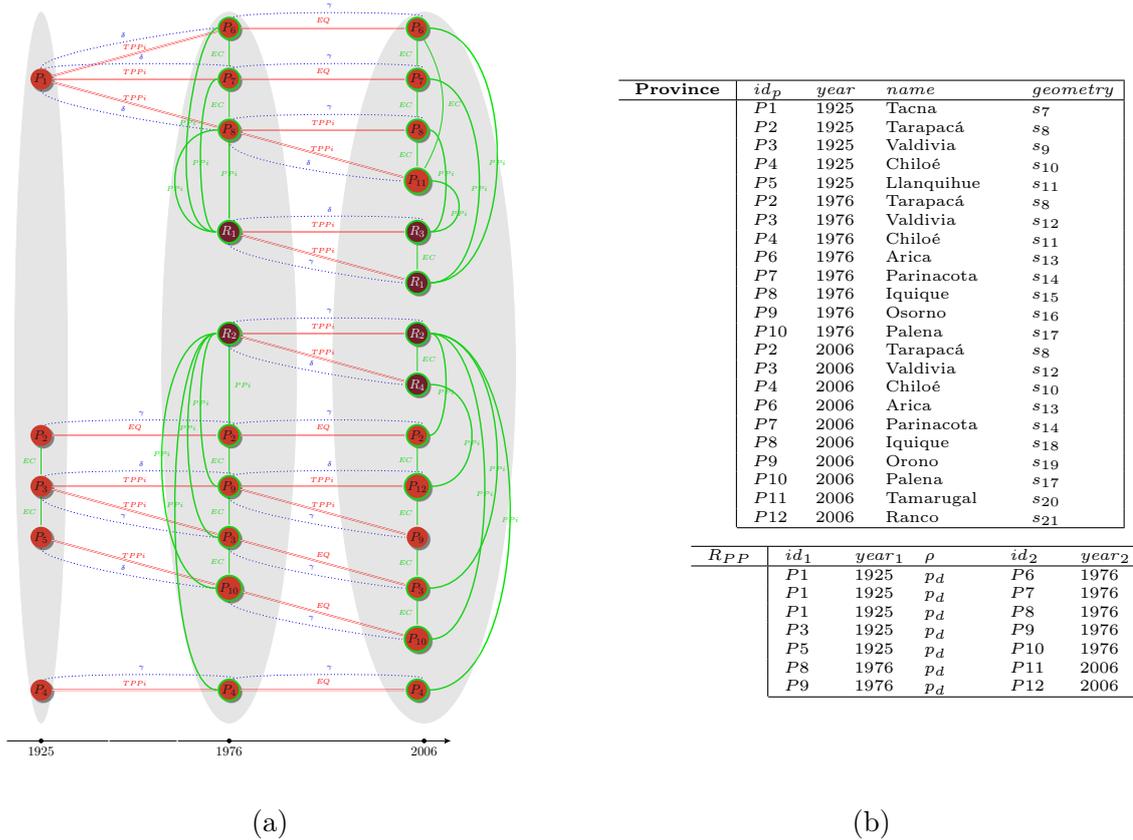


Figure 57 – (a) Le graphe spatio-temporel résultat de la jointure entre régions et provinces ; (b) L'instanciation du graphe des provinces comme schéma de données Σ . L'instanciation du graphe résultat de la jointure dans Σ correspond à l'union du schéma de données des provinces (ci-dessus en b) et celui des régions (Figure 56).

III.3 Spécification des contraintes

III.3.1 Les contraintes inhérentes au modèle : contraintes de domaine (DC)

Dans les bases de données il existe un nombre important de contraintes de domaine relatives à la géométrie des objets spatiaux [Open GIS Consortium Inc., 1999] (e.g. définition d'un polygone, spécification des relations topologiques entre polygones). Dans le cadre de cette thèse, on ne

traitera pas de ces contraintes car pour nous leur sens est indépendant du modèle spatio-temporel. Concrètement, on ne vérifie pas qu'une relation topologique existe ou non, car cette relation provient d'un calcul qui est réalisé sur la géométrie des entités et qui, par conséquent, est indépendant du modèle de données. Nous nous concentrons donc sur les contraintes de domaine relatives aux relations de filiation.

Soit $G_{ST} = (V, E)$ le schéma de données d'un graphe spatio-temporel, avec $O, O' \in V$ et $R \in E$. Les contraintes qui suivent sont associées à la représentation des prédicats relatifs aux entités et aux filiations. Les conséquences des contraintes d'équations III.1 et III.2 sont plutôt relatives aux entités⁷ (on les nomme **RD**, Relations de dépendances), la contrainte d'équation III.3 porte spécifiquement sur la relation de dérivation (on l'appelle **FC**, Contrainte de Filiation), les contraintes d'équations III.4 et III.5 portent sur les dépendances entre les relations de dérivation et de continuation (on les appelle contraintes de dépendance existentielle liée à l'identité (**IED**)). Ces équations III.4 et III.5 correspondent aux contraintes que nous avons posées au chapitre II Figure 28b. et Figure 28a. respectivement.

Relations de dépendances (RD)

Transcription de la relation de continuation, si les identifiants de deux entités sont les mêmes, alors leurs géométries et leurs éventuels attributs aussi :

$$\bar{\forall} (O(id_1, t_1, \bar{u}_1, g_1) \wedge O(id_1, t_2, \bar{u}_2, g_2) \rightarrow (\bar{u}_1 = \bar{u}_2) \wedge (g_1 = g_2)) \quad (\text{III.1})$$

où $\bar{u}_1 = \bar{u}_2$ signifie l'égalité des séquences d'attributs, et $g_1 = g_2$ l'égalité de la géométrie des deux entités.

Conséquence de l'existence d'une relation (de filiation), les deux entités en relation doivent être de la forme :

$$\forall id_1, id_2, t_1, t_2, \delta (R(id_1, t_1, \delta, id_2, t_2) \rightarrow \exists \bar{u}_1, g_1, \bar{u}_2, g_2 (O(id_1, t_1, \bar{u}_1, g_1) \wedge O'(id_2, t_2, \bar{u}_2, g_2))) \quad (\text{III.2})$$

7. La contrainte d'équation III.1 dépend particulièrement de l'application, ici on considère que les attributs et la géométrie d'une entité ne varie pas s'il s'agit de la même entité à un temps différent. Ce n'est pas le cas lorsque par exemple l'on traite d'individus, dont les attributs comme l'âge peuvent varier même s'il s'agit de la même entité.

Contrainte de Filiation (FC)

Définition de la relation de dérivation, si deux entités sont reliées par une relation de dérivation entre deux temps t_1 et t_2 alors leurs identités sont différentes et $t_1 \leq t_2$:

$$\bar{\forall} (R(id_1, t_1, \delta, id_2, t_2) \rightarrow (id_1 \neq id_2) \wedge (t_1 \leq t_2)) \quad (III.3)$$

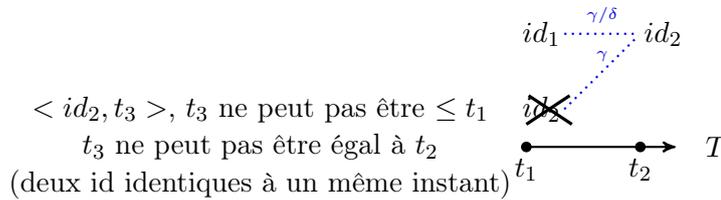
Contraintes de dépendance existentielle liée à l'identité (IED)

Si une entité (à t_2) est issue d'une relation de filiation avec une autre (à t_1), cette entité ne peut entretenir de relation de continuation avec une troisième (à $t_3 \leq t_1$)

$$\bar{\forall} (O_1(id_1, t_1, \bar{u}_1, g_1) \wedge O_2(id_2, t_2, \bar{u}_2, g_2) \wedge O_2(id_2, t_3, \bar{u}_3, g_3) \wedge R_1(id_1, t_1, \rho_f, id_2, t_2) \rightarrow (t_3 > t_2)) \quad (III.4)$$

avec $\rho_f \in \Phi$.

Schéma explicatif de la contrainte III.4 :

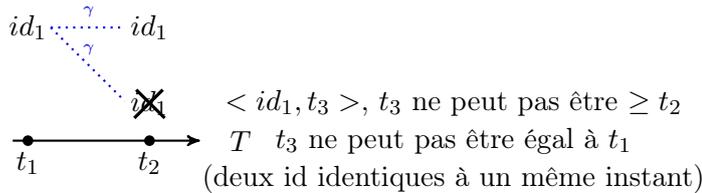


Si une entité (à t_1) entretient une relation de continuation avec une autre (à t_2), elle ne peut continuer son existence par ailleurs (à $t_3 \geq t_1$) :

$$\bar{\forall} (O_1(id_1, t_1, \bar{u}_1, g_1) \wedge O_1(id_1, t_2, \bar{u}_2, g_2) \wedge O_1(id_1, t_3, \bar{u}_3, g_3) \wedge R_1(id_1, t_1, \rho_f, id_2, t_2) \rightarrow (t_3 < t_1)) \quad (III.5)$$

avec $\rho_f \in \Phi$.

Schéma explicatif de la contrainte III.5 :

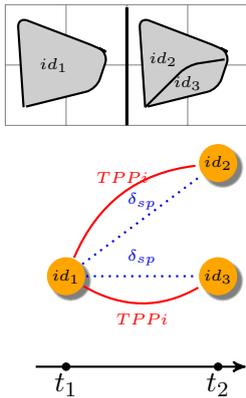


III.3.2 Les contraintes dépendantes d'une application : contraintes sémantiques (SC)

Le nombre de contraintes sémantiques est potentiellement infini puisqu'il est relatif à autant d'applications que l'on peut vouloir modéliser. Nous nous concentrons sur la spécification de contraintes sémantiques relatives à des processus étudiés dans le domaine de la modélisation spatio-temporelle [Claramunt et Thériault, 1995, Galton, 2004]. L'entité est une région de l'espace (entité statique), et l'on peut décrire informellement :

- Division : une région est divisée en deux nouvelles régions.
- Séparation : une partie d'une région se sépare pour créer une nouvelle région.
- Fusion : deux régions se regroupent (et disparaissent) pour former une seule nouvelle région.
- Annexation : une région devient une partie d'une autre (et donc la première disparaît).

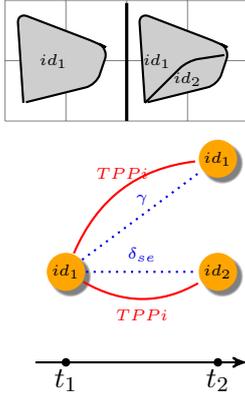
Soient $O_1, O_2, O_3 \in V$ les prédicats relatifs aux entités spatio-temporelles, $R_1, R_2 \in E$ les prédicats relatifs à la relation de filiation du schéma de données du graphe spatio-temporel $G_{ST} = (V, E)$. L'ensemble des valeurs que peut prendre la relation de filiation est $\Phi = \{\gamma, \delta_{sp}$ (division), δ_{se} (séparation), δ_{me} (fusion), δ_{an} (annexion)\}. Les filiations de type de dérivation réfèrent à la sémantique spécifique de la relation de filiation pour chacun de ces processus. Nous présentons Figures 58, 59, 60 et 61 les contraintes qui expriment la sémantique des différents processus spatio-temporels. Dans ces exemples, $\forall i, j \in \mathbb{N}$, l'union géométrique $GeomUnion(g_i, g_j)$ renvoie l'aire égale à l'union des aires de g_i et g_j .



$$\bar{\forall}(O_1(id_1, t_1, \bar{u}_1, g_1) \wedge O_2(id_2, t_2, \bar{u}_2, g_2) \wedge O_3(id_3, t_2, \bar{u}_3, g_3) \wedge R_1(id_1, t_1, \delta_{sp}, id_2, t_2) \wedge R_2(id_1, t_1, \delta_{sp}, id_3, t_2) \rightarrow EQ(g_1, GeomUnion(g_2, g_3))).$$

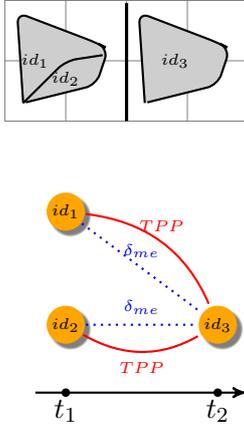
Figure 58 – Division de régions

Nous allons maintenant expliquer et généraliser les différents types de contraintes sémantiques établies sur les processus spatio-temporels précédemment évoqués.



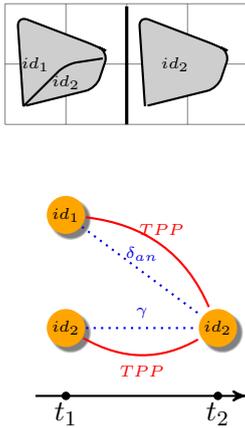
$$\bar{\forall} O_1(id_1, t_1, \bar{u}_1, g_1) \wedge O_1(id_1, t_2, \bar{u}_3, g_2) \wedge O_2(id_2, t_2, \bar{u}_2, g_3) \wedge R_1(id_1, t_1, \delta_{se}, id_2, t_2) \rightarrow \text{EQ}(g_1, \text{GeomUnion}(g_2, g_3)).$$

Figure 59 – Séparation de régions.



$$\bar{\forall} (O_1(id_1, t_1, \bar{u}_1, g_1) \wedge O_2(id_2, t_1, \bar{u}_2, g_2) \wedge O_3(id_3, t_2, \bar{u}_3, g_3) \wedge R_1(id_1, t_1, \delta_{me}, id_3, t_2) \wedge R_2(id_2, t_1, \delta_{me}, id_3, t_2) \rightarrow \text{EQ}(g_3, \text{GeomUnion}(g_1, g_2)).$$

Figure 60 – Fusion de régions.



$$\bar{\forall} (O_1(id_1, t_1, \bar{u}_1, g_1) \wedge O_2(id_2, t_1, \bar{u}_2, g_2) \wedge O_2(id_2, t_2, \bar{u}_3, g_3) \wedge R_1(id_1, t_1, \delta_{an}, id_2, t_2) \rightarrow \text{EQ}(g_3, \text{GeomUnion}(g_1, g_2)).$$

Figure 61 – Annexation de régions.

Soit $\forall i \in \mathbb{N}$, $O_i \in V$ les prédicats relatifs aux entités, et $R_i \in E$ les prédicats relatifs aux relations de filiations du schéma de données du graphe spatio-temporel $G_{ST} = (V, E)$. Nous avons classé les contraintes sémantiques en deux catégories, les Dépendances de Filiation liées à la Topologie

(TFD) ou les Dépendances de Filiation liées à la topologie et à la Géométrie (GFD) :

Dépendance de la Topologie (TFD) liée à la Filiation

Un type de relation de filiation peut contraindre les relations topologiques qui s'opèrent sur la géométrie des entités, d'où la forme générale de contrainte :

$$\bar{\forall}(O_1(id_1, t_1, \bar{u}_1, g_1) \wedge O_2(id_2, t_2, \bar{u}_2, g_2) \wedge R_1(id_1, t_1, \rho_f, id_2, t_2) \rightarrow \Upsilon(g_1, g_2)) \quad (\text{III.6})$$

où $\Upsilon \in \Gamma$ et $\rho_f \in \Phi$.

Dépendance de Filiation liée à la Topologie et à la Géométrie (GFD)

Un type de relation de filiation peut contraindre les relations topologiques qui s'opèrent sur la géométrie des entités, et nécessite l'utilisation d'opérateur pour préciser cette géométrie :

Il en existe plusieurs versions qui dépendent du processus traité (i.e. (i) division, (ii) séparation, (iii) fusion, (iv) annexation) :

$$(i) \quad \bar{\forall}(O_1(id_1, t_1, \bar{u}_1, g_1) \wedge O_2(id_2, t_2, \bar{u}_2, g_2) \wedge O_3(id_3, t_2, \bar{u}_3, g_3) \wedge R_1(id_1, t_1, \rho_f, id_2, t_2) \wedge R_2(id_1, t_1, \rho_f, id_3, t_2) \rightarrow \Upsilon(g_1, \theta(g_2, g_3))).$$

$$(ii) \quad \bar{\forall}(O_1(id_1, t_1, \bar{u}_1, g_1) \wedge O_2(id_2, t_2, \bar{u}_2, g_2) \wedge O_1(id_1, t_2, \bar{u}_3, g_3) \wedge R_1(id_1, t_1, \rho_f, id_2, t_2) \rightarrow \Upsilon(g_1, \theta(g_2, g_3))).$$

$$(iii) \quad \bar{\forall}(O_1(id_1, t_1, \bar{u}_1, g_1) \wedge O_2(id_2, t_1, \bar{u}_2, g_2) \wedge O_3(id_3, t_2, \bar{u}_3, g_3) \wedge R_1(id_1, t_1, \rho_f, id_3, t_2) \wedge R_2(id_2, t_1, \rho_f, id_3, t_2) \rightarrow \Upsilon(g_3, \theta(g_2, g_1))).$$

$$(iv) \quad \bar{\forall}(O_1(id_1, t_1, \bar{u}_1, g_1) \wedge O_2(id_2, t_1, \bar{u}_2, g_2) \wedge O_2(id_2, t_2, \bar{u}_3, g_3) \wedge R_1(id_1, t_1, \rho_f, id_2, t_2) \rightarrow \Upsilon(g_3, \theta(g_2, g_1))).$$

où $\theta \in \Theta$ est un opérateur géométrique qui renvoie une géométrie, $\Upsilon \in \Gamma$, et $\rho_f \in \Phi$.

La différence entre ces deux types de contraintes qui portent sur les relations de filiation, est que certaines imposent une condition définie non seulement par des relations topologiques $\Upsilon \in \Gamma$ mais aussi par des opérateurs géométriques $\theta \in \Theta$ (i.e. les GFD), alors que les autres (i.e. les TFD) ne dépendent que de relations topologiques. Une contrainte (GFD) peut impliquer des contraintes

de type **(TFD)**. Par exemple dans le cas de la division, la composition des géométries des entités résultantes de ce processus doit être égale à la géométrie de l'entité initiale. Il s'agit de la contrainte Figure 58, elle implique que l'on doit avoir $PP(g_2, g_1)$ et $PP(g_3, g_1)$.

III.4 Raisonner et vérifier la consistance

III.4.1 Satisfiabilité de l'ensemble de contraintes

Classiquement, lorsque l'on pose un certain nombre de contraintes d'intégrité ψ , il faut s'assurer que cet ensemble est *satisfiable*. Dire qu'un ensemble est satisfiable signifie qu'il existe une instance non vide D du schéma du graphe spatio-temporel G qui satisfait ψ , ce qui s'écrit $D \models \psi$.

Ici, nous avons choisi de tester la consistance des contraintes de type **IEDs** (Contraintes de dépendance existentielle liée à l'identité), **TFDs** (Contraintes de dépendance de filiation liée à la topologie) et **GFDs** (Contraintes de dépendance de filiation liée à la topologie et à la géométrie), la propriété qu'il faut démontrer est donc la suivante :

Propriété Soit un modèle de donnée Σ du graphe spatio-temporel $G_{ST} = (V, E)$ associé aux contraintes de domaine **RD** et **FC** et un ensemble ψ de contraintes **IEDs**, **TFDs** et **GFDs**, alors ψ est toujours satisfiable.

Nous ne détaillerons pas ici la preuve qui peut être consultée dans l'article [Del Mondo *et al.*, 2011] et en annexe III.

III.4.2 Algorithme de consistance d'un ensemble de contraintes sur une instance de graphe

On considère que les contraintes relationnelles (**RD**) sont satisfaites (équations III.1 et III.2). L'algorithme⁸ 1 permet de vérifier qu'une instance de graphe spatio-temporel est consistante sur les contraintes de domaine (**FC**) et (**IED**) (respectivement les équations III.3 et III.4, III.5) et sémantiques (**TDF**) et (**GFD**) (respectivement les équations III.6 et celles relatives à **GFD**) que

8. La complexité (*dans le pire des cas*) de l'algorithme a été évaluée en $O(mn^2)$ si m est le nombre de contraintes à vérifier et n le nombre de tuples qui composent l'instance D associée au graphe G_{ST} du modèle de données. En effet, l'algorithme est constitué d'une suite de structures de données "Pour" (une pour chaque contrainte à vérifier (m)), et à chaque fois on décrit un double "Pour" sur le nombre de tuples de D (n^2) qui teste chaque relation et les entités relatives à cette relation.

nous avons définies dans les paragraphes précédents. Le passage de cet algorithme sur le modèle de données du graphe spatio-temporel nous garantit qu'il respecte le modèle sur lequel il est construit et la spécificité de l'application modélisée.

Ce chapitre nous a permis d'étudier un volet plus technique du modèle de graphe décrit au chapitre précédent, à savoir le modèle de données. Celui-ci entre davantage dans les détails des différents paramètres du modèle de graphe spatio-temporel. La spécification de contraintes, tant inhérentes au modèle que spécifiques à une application, nous garantit que le modèle de données est cohérent vis à vis de la réalité modélisée (par ces contraintes). Les exemples de contraintes sémantiques que nous avons fournis ouvrent la voie à la modélisation de beaucoup d'autres contraintes sur le même schéma pour d'autres applications. Finalement, via un algorithme de test de la consistance de ces contraintes sur une instance de graphe spatio-temporel, nous proposons une méthode automatique de vérification de consistance des données.

```

Input : Une instance  $D$  du graphe spatio-temporel de schéma  $G_{ST} = (V, E)$  et  $\Sigma = (\mathcal{U}, \mathcal{A}, \mathcal{T}, \mathcal{S}, \mathcal{I}, \mathcal{O}, \Phi, \mathcal{G}, \Gamma, \Theta, \Delta)$ , avec un
ensemble de types de filiation dans  $\Phi$  et les ensembles  $\varphi_{1i}, \varphi_{1ii}$  contiennent les différentes formes de  $IEDs$ , un ensemble  $\varphi_2$  de
 $TFDs$ , les ensembles  $\varphi_{3i}, \varphi_{3ii}, \varphi_{3iii}, \varphi_{3iv}$  pour les différentes formes de  $GTDs$ 
Output : True si  $D$  est consistant vis à vis de la définition des contraintes (FC), (IED), (TDF) et (CTDF) ; sinon False
/* Vérification contrainte (FC) */
foreach  $R(i_1, j_1, \rho_d, i_2, j_2) \in D$  do
  if  $((j_1 > j_2) \text{ or } (i_1 = i_2))$  then
    return false;
  end
end
/* Vérification contraintes (IED) équation III.3 */
foreach  $\psi_1 \in \varphi_{1i}$  do
  Soient  $O_1(id_1, t_1, \bar{u}_1, g_1), O_2(id_2, t_2, \bar{u}_2, g_2), O_1(id_1, t_3, \bar{u}_3, g_3)$  prédicats relatifs aux entités, et  $R_1(id_1, t_1, \rho_d, id_2, t_2)$  prédicat
de filiation dans  $\psi_1$ ;
  foreach  $R_1(i_1, j_1, \rho_d, i_2, j_2) \in D$  do
    foreach  $O_1(i_1, j_3, \bar{c}_3, s_3) \in D$  do
      if  $(j_3 > j_1)$  then
        return false;
      end
    end
  end
end
/* Vérification contraintes (IED) cas équation III.4 */
foreach  $\psi_1 \in \varphi_{1ii}$  do
  Soient  $O_1(id_1, t_1, \bar{u}_1, g_1), O_2(id_2, t_2, \bar{u}_2, g_2), O_2(id_2, t_3, \bar{u}_3, g_3)$  prédicats relatifs aux entités, et  $R_1(id_1, t_1, \rho_d, id_2, t_2)$  prédicat
de filiation dans  $\psi_1$ ;
  foreach  $R_1(i_1, j_1, \rho_d, i_2, j_2) \in D$  do
    foreach  $O_2(i_2, j_3, \bar{c}_3, s_3) \in D$  do
      if  $(j_3 < j_1)$  then
        return false;
      end
    end
  end
end
/* Vérification contrainte (TDF) */
foreach  $\psi_2 \in \varphi_2$  do
  Soient  $O_1(id_1, t_1, \bar{u}_1, g_1), O_2(id_2, t_2, \bar{u}_2, g_2)$  prédicats relatifs aux entités, et  $R_1(id_1, t_1, \rho_f, id_3, t_2)$  prédicat de filiation, et
 $\Upsilon(g_1, g_2)$  relation topologique dans  $\psi_2$ ;
  foreach  $R(i_1, j_1, \rho_f, i_2, j_2) \in D$  do
    Trouver  $O_1(i_1, j_1, \bar{c}_1, s_1), O_2(i_2, j_2, \bar{c}_2, s_2) \in D$ ;
    if  $(\text{not } \Upsilon(s_1, s_2))$  then
      return false;
    end
  end
end
/* Vérification contraintes (GFD) cas (i) */
foreach  $\psi_3 \in \varphi_{3i}$  do
  Soient  $O_1(id_1, t_1, \bar{u}_1, g_1), O_2(id_2, t_2, \bar{u}_2, g_2), O_1(id_1, t_2, \bar{u}_3, g_3)$  des prédicats et  $R(id_1, t_1, \rho_f, id_2, t_2)$  un prédicat relatif à la
relation de filiation, et  $\Upsilon(g_1, \theta(g_2, \dots, g_3))$  une relation topologique dans  $\psi_3$ ;
  foreach  $R(i_1, j_1, \rho_f, i_2, j_2) \in D$  do
    Soient  $O_1(i_1, j_1, \bar{c}_1, s_1), O_2(i_2, j_2, \bar{c}_2, s_2), O_1(i_1, j_2, \bar{c}_3, s_3) \in D$ ;
    if  $(\text{non } \Upsilon(s_1, \theta(s_2, s_3)))$  then
      return false;
    end
  end
end
/* Vérification contraintes (GFD) cas (ii) */
foreach  $\psi_3 \in \varphi_{3ii}$  do
  Soient  $O_1(id_1, t_1, \bar{u}_1, g_1), O_2(id_2, t_1, \bar{u}_2, g_2), O_1(id_1, t_2, \bar{u}_3, g_3)$  prédicats relatifs aux entités,  $R(id_1, t_1, \rho_f, id_2, t_2)$  prédicat de
filiation et  $\Upsilon(g_3, \theta(g_2, \dots, g_1))$  relation topologique dans  $\psi_3$ ;
  foreach  $R(i_1, j_1, \rho_f, i_2, j_2) \in D$  do
    Soient  $O_1(i_1, j_1, \bar{c}_1, s_1), O_2(i_2, j_1, \bar{c}_2, s_2), O_1(i_1, j_2, \bar{c}_3, s_3) \in D$ ;
    if  $(\text{non } \Upsilon(s_3, \theta(s_2, s_1)))$  then
      return false;
    end
  end
end
/* Vérification contraintes (GFD) cas (iii) */
foreach  $\psi_3 \in \varphi_{3iii}$  do
  Soient  $O_1(id_1, t_1, \bar{u}_1, g_1), O_2(id_2, t_2, \bar{u}_2, g_2), O_3(id_3, t_2, \bar{u}_3, g_3)$  prédicats relatifs aux entités,
 $R_1(id_1, t_1, \rho_f, id_2, t_2), R_2(id_1, t_1, \rho_f, id_3, t_2)$  prédicats de filiation, et  $\Upsilon(g_1, \theta(g_2, \dots, g_3))$  relation topologique dans  $\psi_3$ ;
  foreach  $R_1(i_1, j_1, \rho_f, i_2, j_2), R_2(i_1, j_1, \rho_f, i_3, j_2) \in D$  do
    Trouver  $O_1(i_1, j_1, \bar{c}_1, s_1), O_2(i_2, j_2, \bar{c}_2, s_2), O_3(i_3, j_2, \bar{c}_3, s_3) \in D$ ;
    if  $(\text{non } \Upsilon(s_1, \theta(s_2, s_3)))$  then
      return false;
    end
  end
end
/* Vérification contraintes (GFD) cas (iv) */
foreach  $\psi_3 \in \varphi_{3iv}$  do
  Soient  $O_1(id_1, t_1, \bar{u}_1, g_1), O_2(id_2, t_1, \bar{u}_2, g_2), O_3(id_3, t_2, \bar{u}_3, g_3)$  prédicats relatifs aux entités,
 $R_1(id_1, t_1, \rho_f, id_3, t_2), R_2(id_2, t_1, \rho_f, id_3, t_2)$  prédicats de filiation, et  $\Upsilon(g_3, \theta(g_1, \dots, g_2))$  relation topologique dans  $\psi_3$ ;
  foreach  $R_1(i_1, j_1, \rho_f, i_3, j_2), R_2(i_2, j_1, \rho_f, i_3, j_2) \in D$  do
    Trouver  $O_1(i_1, j_1, \bar{c}_1, s_1), O_2(i_2, j_1, \bar{c}_2, s_2), O_3(i_3, j_2, \bar{c}_3, s_3) \in D$ ;
    if  $(\text{not } \Upsilon(s_3, \theta(s_1, s_2)))$  then
      return false;
    end
  end
end
return true;

```

Algorithm 1 : Consistance sémantique sur une instance de graphe

Prise en compte de la granularité spatiale

Sommaire

IV.1 Motivations	91
IV.2 La granularité spatiale dans le modèle de graphe spatio-temporel . . .	95
IV.3 La piste des bigraphes	101
IV.3.1 Définition d'un bigraphe	102
IV.3.2 Application à la problématique spatio-temporelle : dynamique dans les bigraphes	106

IV.1 Motivations

La représentation informatique d'un phénomène implique une description plus ou moins simplifiée de la réalité, et la modélisation et l'analyse sont fonction du niveau de détail disponible et/ou souhaité de cette description. La référence à la notion de *niveau de détail* nous renvoie directement au concept de granularité qui correspond à la vision (et donc la modélisation) d'un phénomène selon différents niveaux de détail [Meentemeyer, 1989]. Ces niveaux sont organisés selon une structure hiérarchique qui permet d'augmenter ou de réduire le détail de la représentation [Frank et Timpf, 1994], dans la limite imposée par les données. Concrètement on ne peut pas accéder à un niveau de détail qui n'est pas présent dans les données.

La possibilité de faire varier la granularité au sein d'un modèle est essentielle si l'on souhaite étudier un même phénomène à plusieurs niveaux de détail. La variation du niveau de granularité spatiale permet des analyses différentes et potentiellement complémentaires d'un même phénomène

[Claramunt et Jiang, 2000]. Par ailleurs, selon le point de vue de l'application ou de l'observateur, le passage à un niveau d'abstraction moins détaillé (généralisation spatiale ou temporelle) peut induire une perte d'information contenue dans les niveaux les plus détaillés. On peut bien sûr préciser ensuite si l'information est perdue ou non. La granularité est usuellement déclinée selon trois dimensions : sémantique, temporelle et spatiale [Vangenot, 2004].

- La granularité sémantique définit le niveau de détail en fonction du niveau d'abstraction conceptuel souhaité. En prenant l'exemple d'une classification de végétaux (Figure 62), le niveau sémantique le plus fin sélectionne les plantes suivant leur espèce (e.g. ronce, figuier, framboisier...), les mêmes éléments peuvent être représentés sous la forme d'étiquettes plus générales des familles *Rosaceae* et *Moraceae*. Au plus haut niveau d'abstraction, on peut finalement ne visualiser que l'ordre global des *Rosales*, sans distinguer les différentes familles ou espèces d'entités.
- La granularité temporelle permet de représenter un phénomène selon différents pas de temps. Figure 63, si on regarde le déroulement d'un évènement pluvieux en posant une granularité temporelle par heure, on saura à chaque heure s'il pleut ou non. Si l'on modélise l'évènement à l'échelle de la journée on identifie si un évènement pluvieux a été observé, mais on ne peut préciser à quel moment dans la journée.
- La granularité spatiale consiste en la représentation d'un phénomène selon différentes résolutions. Figure 64.1, le relief (ici le mont Seorak) est représenté selon deux niveaux de granularité spatiale différents. On considère ici des entités de type polygone classifiées en fonction de la pente et de l'orientation. Dans l'exemple présenté, faire varier la granularité revient à modifier la classification en choisissant des intervalles de valeurs de pente et d'orientation différents. Le niveau de détail est d'autant plus élevé que le nombre de classes est important, ce qui correspond à la distinction de polygone pour des différences très faibles de ces valeurs. Le passage vers un niveau de détail plus faible suppose une réduction du nombre de classes, ce qui entraîne le regroupement des polygones voisins si leurs valeurs sont proches.

En référence à l'approche décrite dans cette thèse, nous nous focalisons sur la définition de la granularité spatiale. L'approche consiste à utiliser les relations d'inclusion (spatiales) entre entités (i.e. recouvrement total ou partiel, relations topologiques de type *O* Figure 3.b). Cependant, les différents niveaux de détail peuvent parfois être interprétés comme sémantiques suivant le type d'application. Par exemple, Figure 64.2, l'objectif est d'étudier les relations entre des personnes et leur espace de vie à trois niveaux de granularité : (1) le niveau limité aux personnes elles-mêmes, sans lien particulier avec l'espace qui les entoure

(e.g. relation de filiation), (2) le niveau relatif aux appartements, (3) le niveau les immeubles. Il y a bien un lien spatial d'inclusion entre ces trois niveaux de granularité, et c'est sur ce critère que l'on différencie ces niveaux, et non pas sur le critère sémantique lié à la nature des entités (e.g. une personne habite un appartement, un immeuble est composé d'appartements). Par la combinaison des niveaux de granularité, on cherche à établir d'éventuels liens entre les personnes et leur lieux d'habitation (e.g. on peut par dérivation retrouver les liens de voisinage). A des niveaux plus macroscopiques, on peut imaginer que les immeubles se trouvent dans des villes, ces villes sont contenues dans des provinces qui elles-mêmes appartiennent à des régions. Dans ce cas, on peut replacer cet exemple dans le contexte de l'application sur les régions et provinces au Chili du chapitre II.

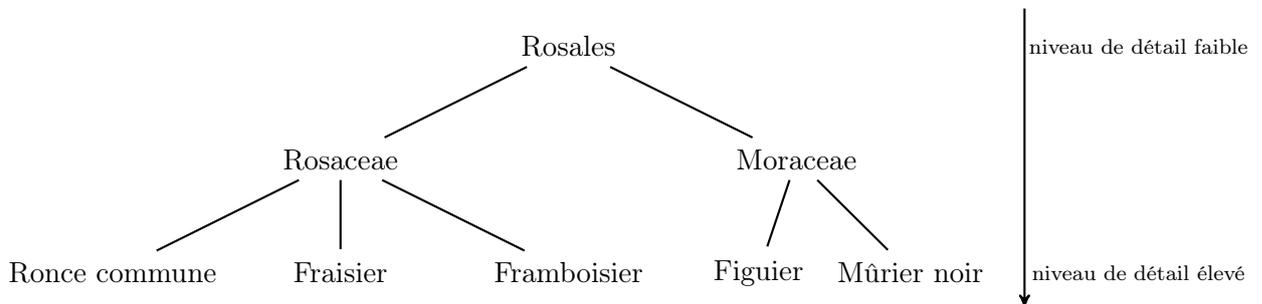


Figure 62 – Trois niveaux de granularité sémantique d'un extrait de la classification végétale de l'ordre des *Rosales*

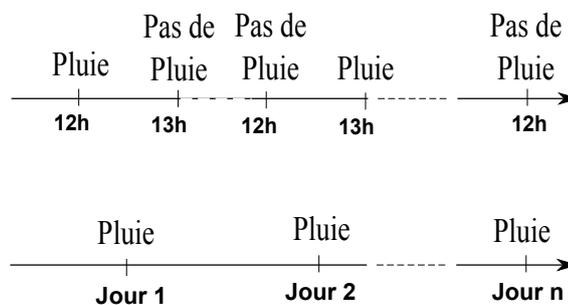


Figure 63 – Deux niveaux de granularité temporelle

Quelle que soit la dimension considérée, le choix d'une granularité optimale pour un phénomène doit permettre de stocker un nombre minimal de données, adapté au niveau de détail dont on a besoin pour son analyse [Camossi, 2004]. Si l'on choisit une granularité trop grossière on risque de ne pas disposer de suffisamment d'information pour répondre au problème modélisé ; si au contraire l'on choisit un niveau de détail trop fin, le nombre d'informations à traiter risque d'être inutilement très important. Le choix du niveau de détail étant lié au type d'analyse, la possibilité de faire plusieurs

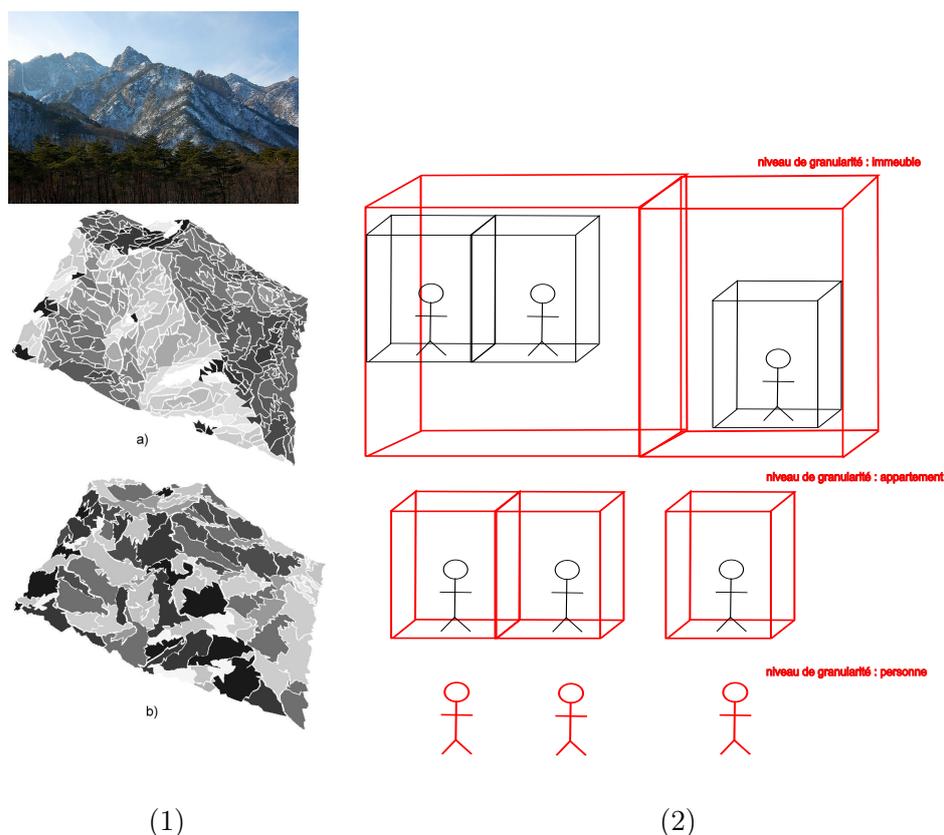


Figure 64 – Deux définitions de la granularité spatiale - (1) Deux niveaux de granularité spatiale en fonction de la pente et de l’orientation pour le mont Seorak (Figure issue de [Saux, 2010]) : le critère de pente est pris dans un intervalle de valeurs plus grand en (b) qu’en (a) ; (2) Trois niveaux de granularité spatiale (et éventuellement sémantique) pour l’étude d’une population : personnes / appartements / immeubles

types d’analyse pour un même phénomène nécessite la mise en place d’un modèle permettant de faire varier la granularité.

Nous décrivons ici un chapitre de prospective dans lequel nous montrons dans une première partie, comment l’intégration de la granularité spatiale dans le modèle de graphe que nous avons précédemment défini au chapitre II est possible. Dans une seconde partie, nous présentons une étude préliminaire¹ traitant d’une structure de graphe plus adaptée au problème de la granularité spatiale : le bigraphe.

Nous avons deux objectifs principaux : (1) Représenter un même phénomène suivant différents niveaux de granularité ; (2) Combiner ces niveaux de granularité afin de déduire des informations supplémentaires sur le phénomène. Pour illustrer le deuxième point, prenons l’exemple des personnes dans des appartements, un exemple de combinaison simple de ces niveaux de granularité peut nous

1. Cette étude est relative aux travaux publiés dans [Stell *et al.*, 2011].

informer sur le fait que deux personnes sont voisines si leurs appartements se touchent.

IV.2 La granularité spatiale dans le modèle de graphe spatio-temporel

Le modèle de graphe spatio-temporel est générique, mais pour présenter une solution d'intégration de la granularité spatiale nous nous appuyerons sur un exemple. La problématique support à notre réflexion concerne la modélisation des relations spatio-temporelles qui existent entre des personnes et leur espace de vie. La variation de la granularité spatiale dans cet exemple, qui a déjà été illustré Figure 64.2, permet d'étudier les relations entre personnes en tant que telles (et les relations de filiations potentielles qui les lient), ou leur localisation dans le bâti (i.e. appartements, immeubles). Sur trois temps t_1, t_2, t_3 , l'idée est de représenter le scénario suivant (Figure 65) en distinguant les relations spatiales des relations de filiation :

Temps t_1

Relations spatiales :

- Les personnes A, B, E_1, E_2 habitent dans un même appartement Δ_{AB} .
- La personne C habite dans un appartement Δ_C .
- Les personnes D, F habitent dans un appartement Δ_{DF} .
- Les appartements Δ_{AB} et Δ_C se trouvent dans l'immeuble I_1 .
- L'appartement Δ_{DF} se trouve dans l'immeuble I_2 .

Relations de filiations :

- A et B sont issus de deux familles distinctes.
- E_1 et E_2 sont les enfants de A et B .
- C et F n'ont pas de lien de filiation.
- D possède un lien de filiation avec A (frère/soeur).

Temps t_2

Relations spatiales :

- E_3 habite dans le même appartement Δ_{DF} que D et F .

Relations de filiations :

- Naissance de E_3 enfant de D et F .
- B décède.

Temps t_3

Relations spatiales :

- Les appartements Δ_{DF} et Δ_{CE_2} sont voisins (dans l'immeuble I_2).
- C déménage dans l'appartement Δ_{CE_2} .
- E_2 déménage dans l'appartement Δ_{CE_2} .

Relations de filiations :

- Pas de changement.

Ce scénario est modélisé par le graphe spatio-temporel de la Figure 66. Par souci de lisibilité, certaines relations ne sont pas matérialisées sur le schéma, mais elles sont censées exister dans les données : les personnes se trouvant dans un même appartement sont considérées en relation spatiale de connexion et deux entités sont en relation de continuation si elles ont le même identifiant. Nous n'avons pas indiqué non plus les relations spatio-temporelles entre personnes, ni entre personnes et appartement/immeuble. Nous n'avons pas indiqué non plus les relations spatiales entre personne et immeuble. Toutes les relations topologiques sont de type O sauf entre les appartements Δ_{DF} et Δ_{E_2C} qui sont reliés par la relation topologique EC . On peut remarquer sur cet exemple que la spécification des relations spatio-temporelles pour une entité statique (e.g. appartement, immeuble) n'est a priori pas nécessaire². En effet dans ce cas, toutes les versions³ d'une entité sont reliées par des relations spatio-temporelles; ces relations, dans le cas d'entités statiques, pourraient donc devenir implicites.

D'un point de vue général, dès lors que le graphe spatio-temporel est construit, on considère que tout l'aspect modélisation du problème a été effectué : en particulier, le choix de la sélection des entités et les niveaux de détails associés. Le choix des entités et des niveaux de granularité dépendent de la dimension spatiale et du problème d'un point de vue sémantique. Sur un tel graphe, certaines questions se posent dont : (1) Comment faire varier la granularité spatiale et représenter les différents niveaux de détail ? (2) Peut-on résoudre des questions relatives à une combinaison d'informations contenues dans plusieurs niveaux de détail ?

2. Sauf si cette entité statique est déformable (e.g. agrandissement d'un immeuble).

3. On rappelle la définition proposée en bas de page 27 : J'appelle *version* d'une entité, une entité définie à un instant t : une entité possède autant de version(s) que d'instant(s) où elle est présente.

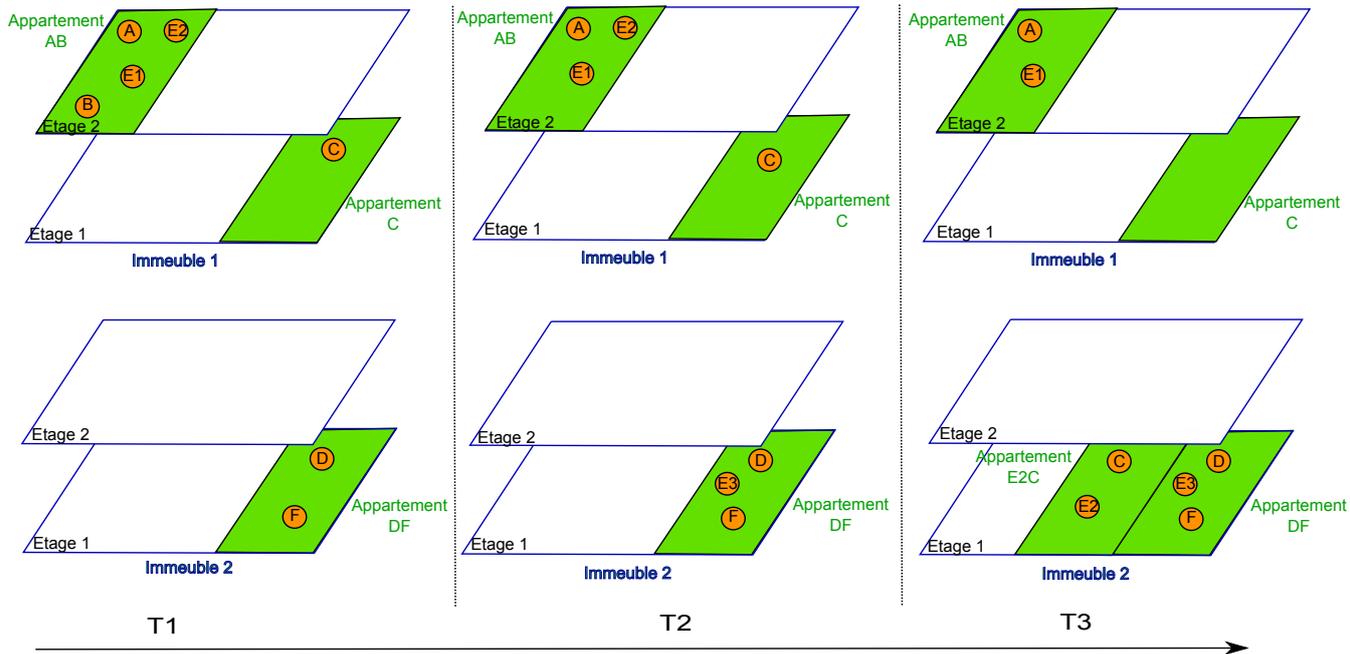


Figure 65 – Scénario de l'évolution de personnes dans leur lieux de vie, sur le domaine de temps $T = \{T1, T2, T3\}$. Le bâti se compose de deux immeubles à deux étages, chaque étage comporte trois appartements.

(1) **Faire varier la granularité spatiale et représenter les niveaux de détail** : Faire varier le niveau de détail de la représentation d'un phénomène c'est finalement sélectionner les données qui répondent à ce niveau de détail. Sur un graphe spatio-temporel, la sélection d'entités peut se faire assez facilement. Par ailleurs, le seul moyen que nous avons de représenter l'inclusion spatiale (i.e. la granularité) dans un graphe spatio-temporel est une arête représentant une relation d'inclusion spatiale (e.g. PP) entre deux entités.

La complexité de la représentation proposée Figure 66 nous suggère que l'étude des différents niveaux de détail ne peut se faire que par extraction de ces niveaux. Etant en présence d'un graphe, on peut se demander s'il n'existe pas des moyens issus de travaux relatifs à cette structure qui permettent d'extraire du graphe les niveaux de granularités souhaités. L'intégration de la granularité spatiale suppose de définir un opérateur qui permette le regroupement⁴ de sommets selon des critères spatio-sémantiques. Dans ce contexte, on peut s'inspirer de l'opérateur d'amalgamation spécifié dans [Stell, 1999, Stell et Worboys, 1999] qui peut définir le passage d'un niveau de granularité à un autre. En effet, l'opérateur regroupe des entités pour décrire un niveau de granularité moins élevé, et l'on pourrait définir la fonction inverse qui affine le niveau de granularité. Cet opérateur d'amalgamation

4. Le regroupement de sommets est valable pour la généralisation, l'opérateur inverse est bien sur nécessaire lorsqu'il s'agit de passer à un niveau de détail plus fin.

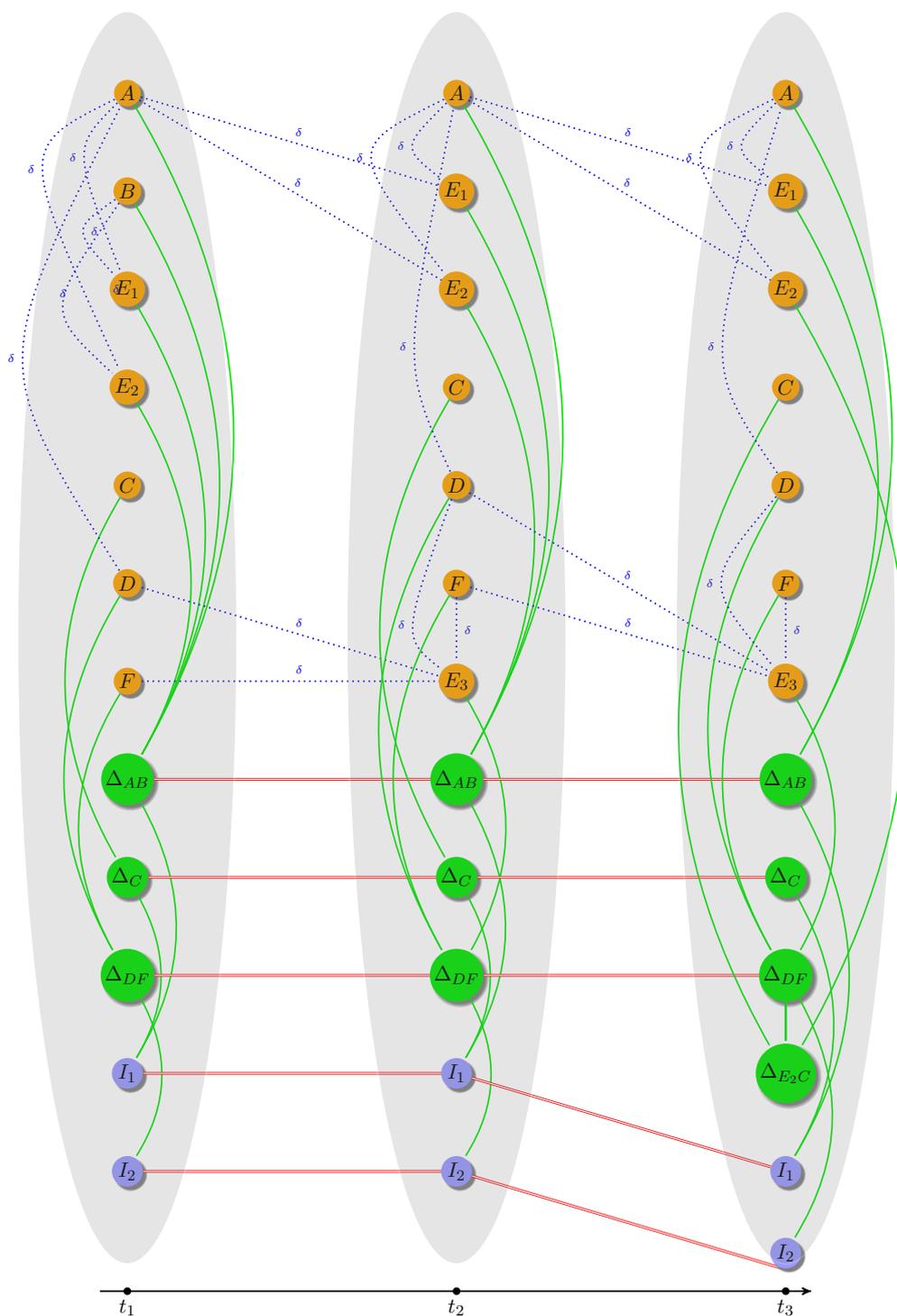


Figure 66 – Modélisation par un graphe spatio-temporel d'un scénario présenté Figure 65 sur plusieurs niveaux de granularité spatiale.

peut être vu comme un opérateur de généralisation spatio-sémantique. Il est générique, et il faut donc le spécifier pour ce type de graphe spatio-temporel et aussi pour chaque application modélisée

par le graphe : décider sur quels critères on change de niveau de granularité, et quelles sont les entités prises en compte pour chaque niveau de granularité. La Figure 67 décrit l'application de cet opérateur sur le temps t_1 de notre exemple.

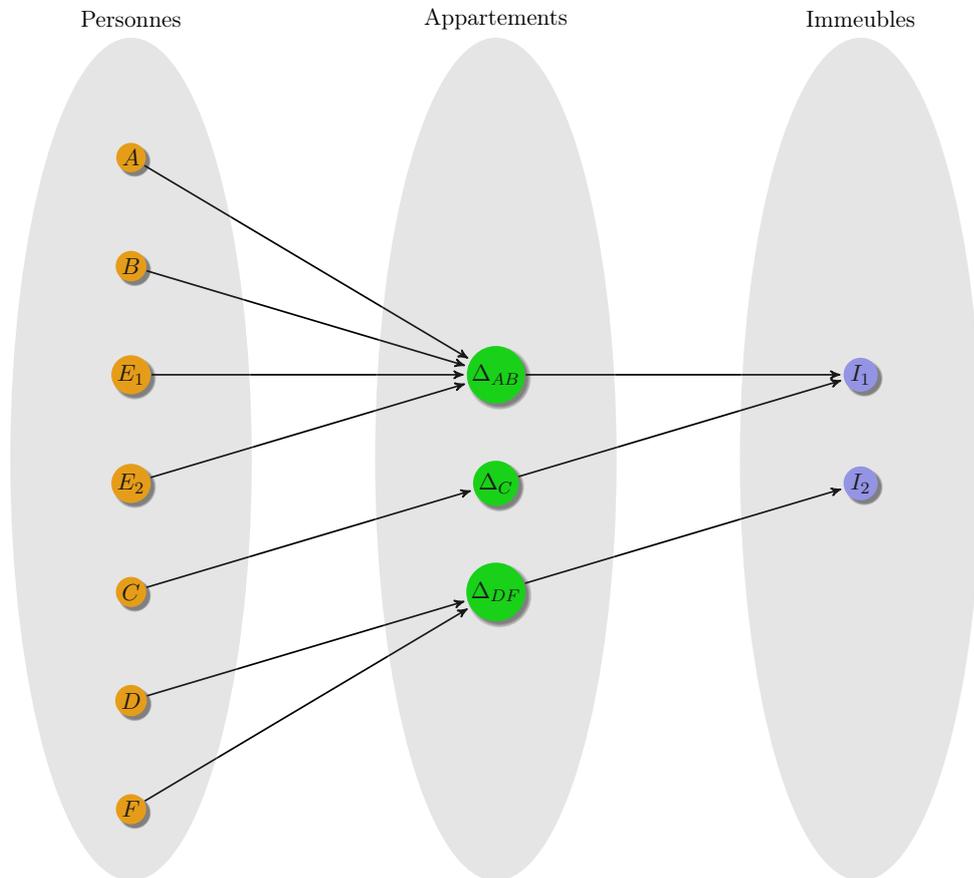


Figure 67 – Opérateur d'amalgamation sur l'exemple au temps t_1 : Après passage de cet opérateur sur le graphe spatio-temporel, un niveau de granularité représente un des trois ensembles d'entités décrits sur ce dessin : les personnes, les appartements ou les immeubles suivant le niveau de détail choisi. Les liens (arcs en noirs sur le dessin) entre ces trois types d'entités se retrouvent par l'utilisation de cet opérateur.

De manière générale, si l'on veut que le graphe spatio-temporel résultat d'une opération sur la granularité (e.g. ici l'amalgamation), soit conforme à la définition de notre modèle de graphe, il doit contenir l'ensemble des entités présentes aux temps considérés (i.e. les entités concernées par le niveau de détail choisi). Il matérialise aussi l'ensemble des relations entre ces entités. On s'abstrait dans un premier temps des relations de filiation afin de se concentrer uniquement sur les relations spatiales.

Il s'agit donc de distinguer les différents niveaux de granularité spatiale et leur hiérarchie (i.e. Ici trois niveaux, personnes, appartements, immeubles). Pour cela, sur l'exemple proposé, on utilise l'opérateur d'amalgamation pour spécifier ces niveaux, et déterminer "la plus longue chaîne

d'inclusion" (i.e. ici, $\text{Personne} \subset \text{Appartement} \subset \text{Immeuble}$ et non pas $\text{Personne} \subset \text{Immeuble}$). On peut ainsi définir l'opérateur d'amalgamation :

$$f_{xy} : X \rightarrow Y \text{ telle que } \forall x \in X, \exists y \in Y \text{ tel que } x \text{ Rel } y, \text{ Rel} \in \{PP, EQ\}.$$

Ce qui dans notre exemple nous permet de conserver les liens, par exemple entre le niveau Appartement et Immeuble :

$$f_{AppIm} : \{\Delta_{AB}, \Delta_C, \Delta_{DF}\} \rightarrow \{I_1, I_2\} \text{ et } \Delta_{AB} \text{ Rel } I_1, \Delta_C \text{ Rel } I_1, \Delta_{DF} \text{ Rel } I_2$$

Dans cet exemple, sélectionner un niveau de granularité revient à sélectionner un type d'entité et les relations associées (i.e. un sous graphe du graphe spatio-temporel initial). Dans le cas général il peut exister des niveaux de granularité spatiale sur une application comprenant plusieurs types d'entité.

En conclusion :

↔ La variation de la granularité sur un graphe spatio-temporel est réalisée par une amalgamation adéquate des données.

↔ Représenter un niveau de granularité revient à extraire un sous graphe du graphe initial. En toute rigueur, il faut néanmoins prévoir de stocker les relations entre les différents niveaux (ici matérialisé grâce à la description informelle de l'opérateur d'amalgamation pour l'application désirée Figure 67).

Ainsi nous pouvons distinguer les différents niveaux de granularité et leur hiérarchie, afin d'associer chaque entité à un niveau de granularité, et permettre l'interrogation des données suivant un certain niveau de granularité.

(2) Répondre à des questions relatives à une combinaison d'informations contenues dans plusieurs niveaux de détail :

Les fonctions relatives à ces questions prennent en argument des entités qui appartiennent à des niveaux de détail différents, et suppose donc de connaître le niveau de détail auquel est liée chaque entité (impliqué par le point précédent (1)). Ensuite, l'interrogation du graphe revient à des requêtes classiques. Prenons l'exemple d'une requête qui prend en arguments l'ensemble des entités personnes et des entités appartements⁵, et qui renvoie *vrai* si les personnes F et C sont dans des appartements voisins. Cette requête peut se décomposer en trois étapes :

- sélectionner les niveaux de granularité relatifs aux personnes et aux appartements pour un temps donné,
- déterminer les appartements dans lesquels sont localisés les personnes F et C (relation

5. Les entités de type personne et de type appartement sont bien de deux niveaux de granularité différents.

- topologique PP),
- tester si ces appartements sont en relation spatiale EC .

En conclusion de cette partie IV.2, nous avons montré comment nous pourrions intégrer la granularité spatiale (i.e. variation du niveau de détail et exemple de requête faisant appel à des entités de niveaux de granularité différents.) dans le modèle de graphe spatio-temporel défini au chapitre II. À partir du constat que la granularité spatiale implique une représentation de plusieurs objets encapsulés, nous avons finalement eu l'opportunité de réfléchir à l'utilisation d'une structure basée sur les bigraphes. Cette structure répond à des exigences spatio-temporelles proches de celles de notre modèle (e.g. représentation des relations spatiales et de filiations) mais où l'intégration de la granularité spatiale est optimisée (e.g. les niveaux de détails peuvent être explicitement représentés).

IV.3 La piste des bigraphes

La représentation d'objets 'encapsulés' sous la forme d'arbre d'inclusion a été abordée dans des travaux comme ceux de [Duckham *et al.*, 2010, Jiang et Worboys, 2009]. L'organisation hiérarchique des différents niveaux de détail dans la représentation de la granularité spatiale nous invite à utiliser ce type de structure. Il se trouve que les arbres d'inclusion sont utilisés dans une structure aux propriétés intéressantes au vu de notre modélisation spatio-temporelle : les bigraphes. Cette modélisation sous forme de bigraphes permet de conserver les liens de filiation entre les entités, de spécifier les modifications spatiales qui s'opèrent entre deux temps, et de représenter et de faire varier le niveau de granularité spatiale. Les relations spatiales entre entités autres que l'inclusion ne sont pas prévues au départ dans la définition des bigraphes, nous proposerons une piste pour parer à ce problème. Le modèle basé sur les bigraphes repose sur la combinaison de deux graphes : un graphe de *liens* (relations de filiation) et un graphe de *places* (relations spatiales). Un exemple de bigraphe est proposé Figure 68, nous détaillerons chaque élément de cette structure dans les paragraphes qui suivent⁶.

6. En complément, toutes les définitions relatives aux bigraphes peuvent être consultées dans le livre de [Milner, 2009]

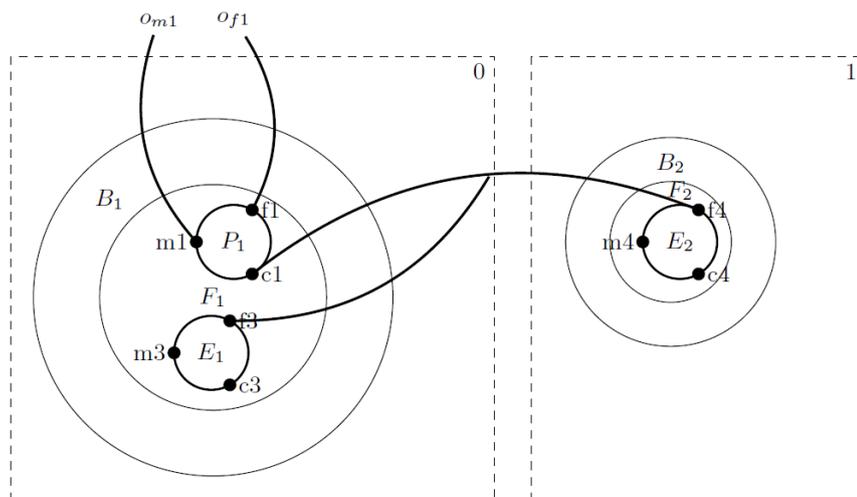


Figure 68 – Exemple de bigraphe : A un instant donné l'espace dans lequel sont localisées les entités est identifié par deux *sites* (0 et 1). On observe par ailleurs deux entités de type *immeuble* (B_1 , B_2), deux entités de type *appartement* (F_1 , F_2) respectivement dans B_1 et B_2 , et trois entités de type *personne* P_1 et E_1 dans F_1 et E_2 dans F_2 .

IV.3.1 Définition d'un bigraphe

Entités et relation de filiation

Une entité (Figure 69) est représentée par un *nœud*. Suivant le type de l'entité, ce nœud possède aucun ou plusieurs *port(s)*. Le choix de ce nombre dépend du concept d'identité et des relations de filiation entre les entités. Par exemple, une personne (Figure 69a) peut posséder trois ports parce qu'elle entretient des liens de filiation avec son père, sa mère et ses enfants potentiels. Cela modélise le fait que l'identité des enfants dépend de celle des parents (Figure 69b). Par contre Figure 69.c, nous avons considéré qu'un immeuble ne possède pas de port car il n'existe pas de relation d'identité avec un autre immeuble ou un autre type d'entité (on peut le voir dans son contexte pour les immeubles B_1 et B_2 de la Figure 68). Ce choix est évidemment arbitraire et dépend du problème posé. Dans cette modélisation la relation de continuation est matérialisée par le fait que les entités possèdent un même identifiant.

Cette représentation nous amène à définir l'hypergraphe appelé *link graph* ou graphe de liens $G = (V, ar, E, link)$ où :

- V est l'ensemble des entités.
- $ar : V \rightarrow \mathbb{O}$ est une fonction, où \mathbb{O} est l'ensemble fini des ordinaux. Cette fonction fait le

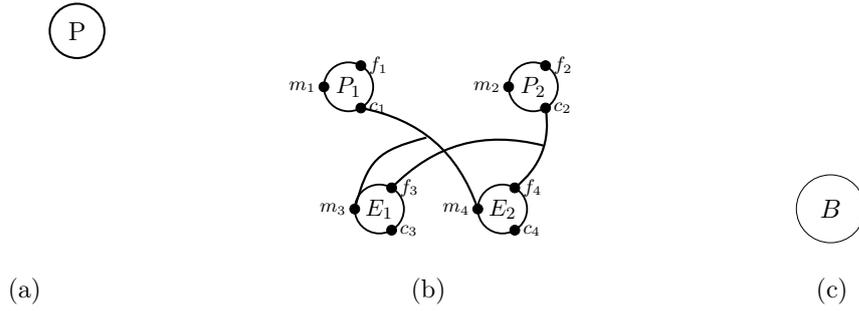


Figure 69 – Exemples de représentation d’entités et leurs liens de filiation : (a) Une personne P possède potentiellement 3 ports relatifs à sa mère, son père et ses enfants ; (b) Deux personnes P_1 et P_2 et leurs enfants E_1 et E_2 ; (c) Un immeuble B qui ici ne possède pas de lien de filiation

lien entre une entité et son type (e.g. personne, appartement, immeuble) défini dans \mathbb{O} .

- E est l’ensemble des (hyper)arêtes de filiation entre les entités (on les appelle aussi *inner links* ou liens intérieurs).
- $link : P_v \rightarrow E$ est une fonction, où P_v est un ensemble de ports. Cette fonction fait la relation entre les arêtes et les ports respectifs sur lesquels elles sont connectées.

Les parties de l’espace où peuvent exister⁷ les entités sont exprimées par le concept de *site* (Figure 70). Chaque site est identifié par une valeur prise dans \mathbb{O} . Si deux entités sont en relation de filiation mais localisées sur deux sites différents dont l’un n’est pas représenté, on conserve la relation de filiation par des liens de filiation "coupés" appelés *outer links* ou liens extérieurs. On définit la fonction $F^L : \mathbb{O} \rightarrow X$ avec X l’ensemble qui contient tous les noms des liens extérieurs.

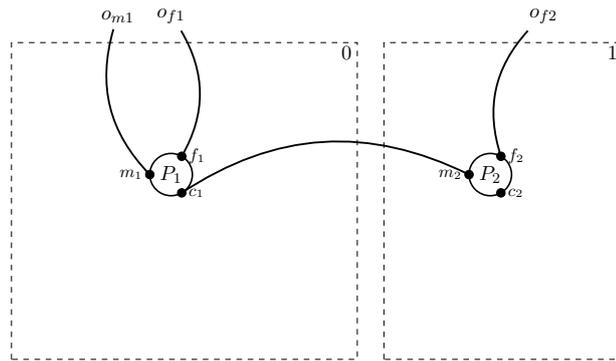


Figure 70 – Exemple de *site* (0 et 1) et de liens extérieurs (o_{m1}, o_{f1}, o_{f2})

Dans l’exemple du graphe de liens Figure 71, $V = \{P_1, P_2, E_1, E_2\}$; et si l’on considère que le type d’entité *Personne* est représenté par '1', $ar(P_1) = ar(P_2) = ar(E_1) = ar(E_2) = 1$; $E = \{ep_1, ep_2\}$; $P_v = \{f_1, m_1, c_1, f_2, m_2, c_2, f_3, m_3, c_3, f_4, m_4, c_4\}$; $link(c_1) = ep_1, link(c_2) = ep_2$,

7. Un site peut être vide.

$link(m_3) = ep_1$, $link(f_3) = ep_2$ etc.

Et $F^L : \emptyset \rightarrow \{O_{f_1}, O_{m_1}, O_{f_2}, O_{m_2}\}$, \emptyset car les entités n'appartiennent à aucun site défini, donc on indique l'ensemble vide par défaut.

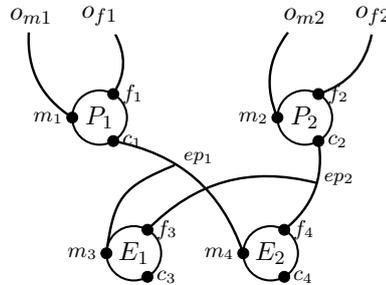


Figure 71 – Exemple de graphe de liens

Entités et relation spatiale

L'unique relation spatiale que peut représenter un bigraphe est la relation d'inclusion. L'inclusion spatiale des entités (qui définit le changement du niveau de détail) est représentée par une *forêt*. Chaque *arbre* de cette forêt décrit quelle entité est incluse dans telle autre. La *racine* de chaque arbre représente le site dans lequel se trouvent les entités considérées.

Ceci nous permet d'introduire le deuxième élément d'un bigraphe : le graphe de places (ou *place graph*). Il est défini par la fonction F^P : Nombre d'interface(s) intérieure(s) \rightarrow Nombre d'interface(s) extérieure(s). Le graphe de place est donc une forêt dont les racines sont des interfaces (i.e. sites) extérieures, et les nœuds des arbres sont soit des entités, soit des interfaces (i.e. sites) intérieures.

Par exemple Figure 72.a, les entités de type personne P_1 et E_1 sont incluses dans l'entité de type appartement F_1 . Cet appartement est lui même inclus dans l'immeuble B_1 , et B_1 est localisé sur un site numéroté 0.

Sur cette même figure, le nombre d'interface(s) extérieure(s) est égal⁸ à $|\{0, 1\}| = 2$ car il y a deux racines dans la forêt (i.e. deux sites extérieurs aux arbres). Le nombre d'interface(s) intérieure(s) est égal à 0 car il n'y a pas de site dans les arbres. Par conséquent $F^P = 0 \rightarrow 2$. A contrario, Figure 72.b, $F^P = 2 \rightarrow 1$ car il existe deux interfaces internes (i.e. sites 1 et 2) et une extérieure (i.e. site 0).

Une alternative possible à la seule modélisation de l'inclusion dans le graphe de places serait

8. Sachant que $|X|$ renvoie le nombre d'éléments contenus dans l'ensemble X .

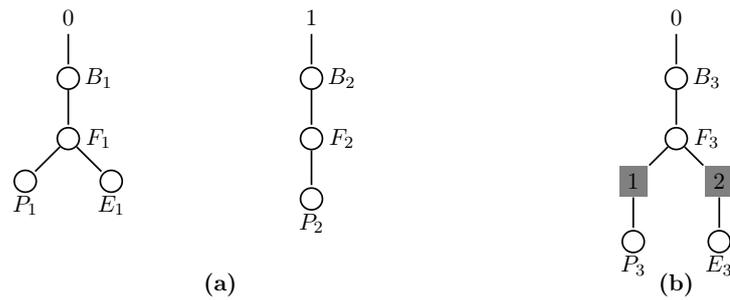


Figure 72 – Exemples de graphes de places : (a) possède deux interfaces extérieures et aucune intérieure ; (b) possède une interface extérieure et deux interfaces intérieures.

de considérer non plus des arbres mais des *DAG* (Graphes orientés acycliques). Un exemple est proposé Figure 73 avec la modélisation du chevauchement (*PO*) des entités *A* et *B*. Un problème potentiel dans le *DAG* Figure 73.b est que les nœuds de l'arbre ne représentent plus seulement une entité mais une relation entre entités, ce qui peut entacher l'homogénéité du graphe. On peut aussi remarquer que si le nombre de relations topologiques à faire figurer est très grand, la représentation du graphe peut rapidement devenir illisible. Cependant, si la visualisation d'un tel graphe n'est pas toujours pertinente, l'exploitation du contenu du *DAG* suite à une implémentation d'un tel modèle, resterait possible.

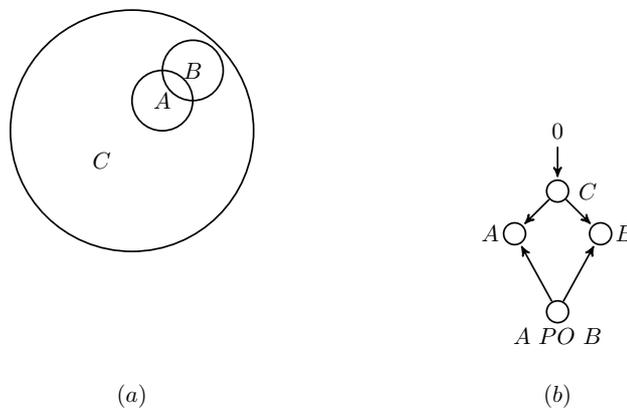


Figure 73 – Représentation des relations spatiales entre entités sous la forme : (a) d'une représentation ensembliste ; (b) d'un graphe orienté acyclique - Permet la modélisation de l'ensemble des relations spatiales du *RCC8* (ici le chevauchement *PO*), a priori la disjonction ne serait pas matérialisée afin de minimiser le nombre de liens.

Modélisation sous forme de bigraphe

L'union des deux graphes définis précédemment, le graphe de liens et le graphe de places, constitue le bigraphe. Un bigraphe F est défini par $F = (F^P, F^L)$ avec F^P et F^L relatifs respectivement

au graphe de places et au graphe de liens. Dans l'exemple Figure 74, le bigraphe est défini par :

$$F^P : 0 \rightarrow 2,$$

$$F^L : \emptyset \rightarrow \{o_{m1}, o_{f1}\},$$

$$\text{Par conséquent } F : (0, \emptyset) \rightarrow (2, \{o_{m1}, o_{f1}\})$$

Cet exemple représente un père P_1 habitant dans un appartement F_1 avec son enfant E_1 . P_1 a un autre enfant E_2 mais qui vit dans un autre appartement F_2 . F_1 et F_2 se trouvent dans deux immeubles distincts B_1 et B_2 qui se trouvent dans des lieux différents, potentiellement éloignés. Les liens extérieurs O_{m1} et O_{f1} signifie que l'on connaît les parents de P_1 mais qu'ils ne se trouvent pas dans un lieu qui nous intéresse d'étudier pour l'instant (i.e. un autre site).

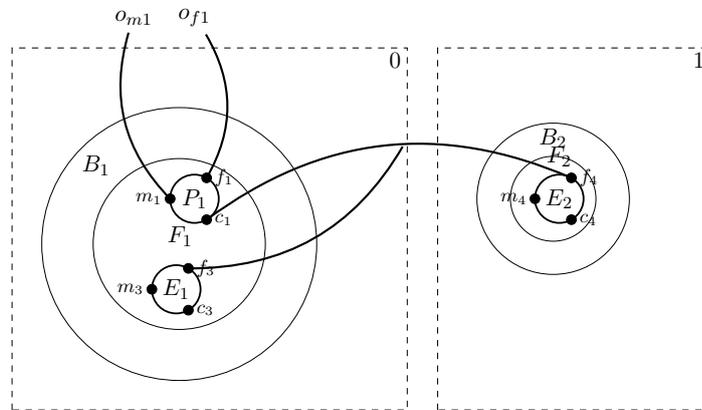


Figure 74 – Exemple de bigraphe

Un exemple d'expression de la granularité spatiale dans un bigraphe est proposé Figure 75. On peut distinguer trois niveaux de granularité : les immeubles, les appartements et les personnes. Le graphe de places relatif à ce bigraphe est décrit Figure 76. Le graphe de liens est similaire à la représentation des relations de filiations relatives au niveau de granularité Personnes, auxquelles il faut ajouter les entités immeubles et appartements, sans aucun lien supplémentaire.

IV.3.2 Application à la problématique spatio-temporelle : dynamique dans les bigraphes

Nous venons de définir le concept de bigraphe à un instant donné. Afin de tendre vers un modèle spatio-temporel capable de modéliser des évolutions, nous allons fournir quelques éléments qui permettent d'introduire la dynamique dans cette structure.

Il n'y a pas de moyen de représenter visuellement la dynamique d'un phénomène dans un

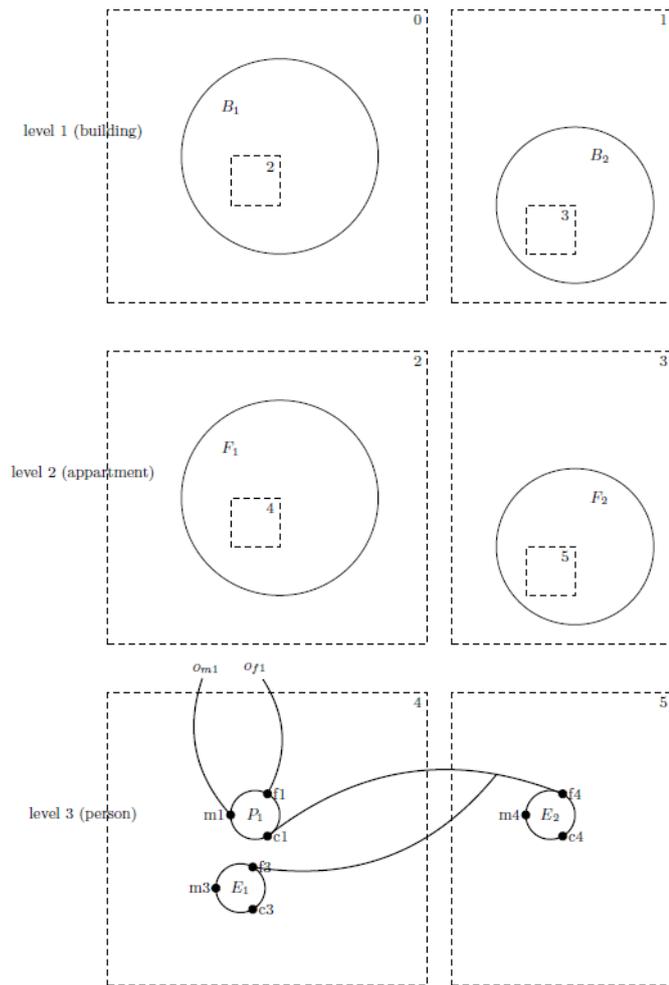


Figure 75 – Un exemple de modélisation des différents niveaux de granularité spatiale avec un bigraphe (équivalent au bigraphe de la Figure 74).

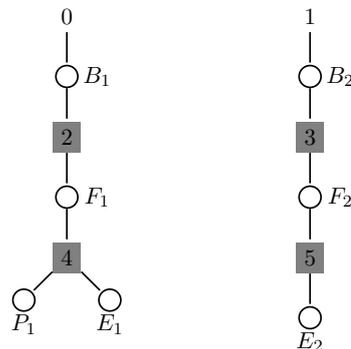


Figure 76 – Graphe de places relatif au bigraphe de la Figure 75. On retrouve les sites extérieurs 0 et 1, les sites intérieurs 2, 3, 4 et 5, ainsi que l'ensemble des entités de la représentation équivalente Figure 74.

bigraphe. Il faut passer par des *règles de dynamique*. Une règle décrit un mouvement, par exemple, le déplacement d'une personne qui rentre (Figure⁹ 77) ou qui sort (Figure 78) d'un immeuble respectivement représentés par les règles :

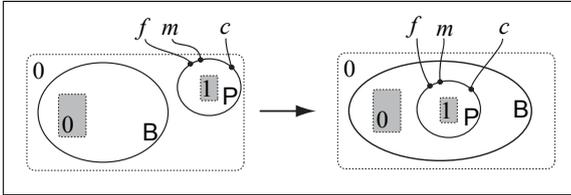


Figure 77 – Représentation de la règle *Entrer* :
 $B.id_0|P_x.id_1 \rightarrow B.(id_0|(P_x.id_1))$

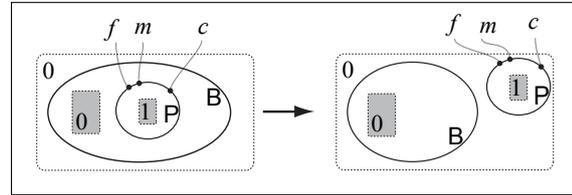


Figure 78 – Représentation de la règle *Sortir* :
 $B.(id_0|(P_x.id_1)) \rightarrow B.id_0|P_x.id_1$

Soit x l'ensemble des ports et/ou noms des liens extérieurs de P (dans l'exemple il n'y a que des noms des liens extérieurs car il n'existe pas de relation de filiation entre un immeuble et une personne).

- *Entrer* : $B.id_0|P_x.id_1 \rightarrow B.(id_0|(P_x.id_1))$
- *Sortir* : $B.(id_0|(P_x.id_1)) \rightarrow B.id_0|P_x.id_1$

Ces deux règles sont construites de la même façon :

Le symbole $.$ dans l'expression $B.id_0$ signifie que le site de numéro 0 **est inclus** dans l'entité B .

Le symbole $|$ dans l'expression $id_0|P_x.id_1$ signifie que le site de numéro 0 et l'entité P (qui contient le site id_1) **n'entretiennent pas** de relation d'inclusion.

L'(es) interface(s) extérieure(s), ici 0, n'est (ne sont) pas représentée(s). Autrement dit, on considère qu'il existe systématiquement un univers qui comprend l'ensemble de la dynamique à modéliser, ici cet univers est assimilé au site extérieur 0.

Pour ces deux règles (i.e. *Entrer*, *Sortir*) nous avons spécifié tous les sites inclus dans les entités, or, s'il n'y a que des sites dans une entité, on peut les omettre de la règle, par exemple la règle *Entrer* devient : $B|P_x \rightarrow B.(id_0|P_x)$.

Si l'on généralise cette approche et que l'on traite d'entités sans type particulier, on peut définir les processus de fusion (Figure 79) et de division (Figure 80) d'entités :

Soient x l'ensemble des ports et/ou noms des liens extérieurs des entités E, F et G et y l'ensemble des ports et/ou noms des liens extérieurs de l'entité H ,

9. La plupart des figures de cette partie sont extraites ou adaptées de [Stell et al., 2011].

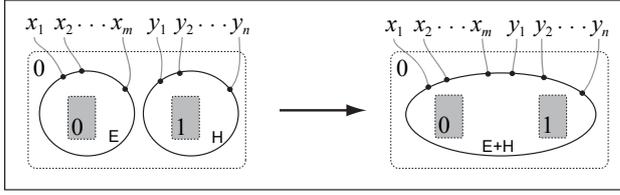


Figure 79 – Représentation de la règle *Fusion* :
 $E_x|H_y \rightarrow (E + H)_{xy}$

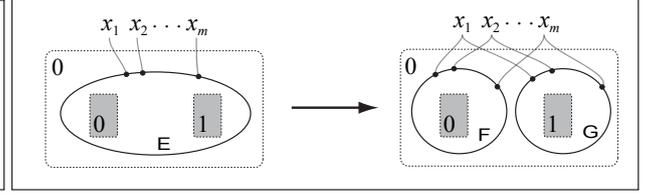


Figure 80 – Représentation de la règle *division* :
 $E_x \rightarrow F_x|G_x$

- *Fusion* : $E_x|H_y \rightarrow (E + H)_{xy}$
- *Division* : $E_x \rightarrow F_x|G_x$

Le symbole $+$ signifiant l'union¹⁰ des deux entités.

Si l'on considère que le nombre de ports d'une entité reflète son type (voir section IV.3.1), alors les définitions des processus précédents impliquent que par fusion la nature d'une entité change, et par division les liens de filiations entre les entités issues de la division perdurent. Cette conséquence étant trop contraignante, il faudrait donc pouvoir relaxer la relation entre le nombre de ports et le type de l'entité en supposant qu'il n'est pas forcément intimement lié.

De la même façon que l'on a défini les processus de fusion et de division d'entité, on s'intéresse à deux autres processus : la *capture* et l'*aspiration*. Ces processus interviennent à partir du moment où l'on se trouve en présence d'entités telles qu'une foule par exemple. Une foule est constituée d'individus, et chaque individu entretient ses propres liens avec le reste de la foule. Dans le cas présent, on ne s'intéresse qu'au type d'entité foule tout en conservant sa propriété de division et les liens entre les différentes parties¹¹ de cette foule. La Figure 81 illustre le cas où une partie de la foule F (i.e. H) est capturée et se retrouve incluse dans la foule G mais en conservant ses liens de filiations avec F . Au contraire dans la Figure 82, la partie de la foule I qui se trouve dans G est fusionnée dans la foule F . On peut remarquer que les deux règles sous-jacentes, *Capture* et *Aspiration* sont des versions légèrement différentes des processus de *Division* ou de *Fusion* mais la sémantique est proche.

Soit x l'ensemble des ports et/ou noms des liens extérieurs des entités F et H , y et z l'ensemble des ports et/ou noms des liens extérieurs respectivement de l'entité G et I ,

- *Capture* : $F_x.(id_0|id_2|G_y.(id_1)) \rightarrow F_x.(id_2|G_y.(id_1|H_x.(id_0)))$
- *Aspiration* : $F_x.(id_2|G_y.(id_1|I_z.(id_0))) \rightarrow F_{xz}.(id_0|id_2|G_y.(id_1))$

10. La sémantique exacte de l'union est définie selon l'application modélisée.

11. Une partie de la foule correspond à un ou plusieurs individus. On considère donc qu'un individu isolé possède le type d'entité "foule"

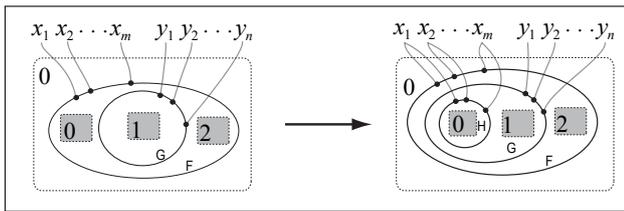


Figure 81 – Représentation de la règle
Capture : $F_x.(id_0|id_2|G_y.(id_1)) \rightarrow$
 $F_x.(id_2|G_y.(id_1|H_x.(id_0)))$

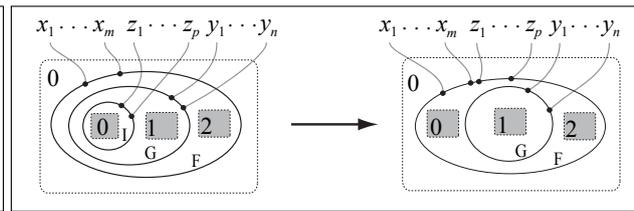


Figure 82 – Représentation de la règle
Aspiration : $F_x.(id_2|G_y.(id_1|I_z.(id_0))) \rightarrow$
 $F_{xz}.(id_0|id_2|G_y.(id_1))$

En conclusion, les bigraphes fournissent une solution de représentation explicite de différents niveaux de granularité spatiale d'un même phénomène. Faire varier cette granularité revient à sélectionner un niveau de granularité. La combinaison des différents niveaux de détail paraît équivalente à celle décrite pour le graphe spatio-temporel, avec lors de l'élaboration de la requête, la sélection des niveaux de détail nécessaires. Bien que l'on puisse disposer des informations concernant les changements spatiaux grâce aux règles de dynamique, on peut néanmoins regretter la non-représentation explicite des liens temporels (spatial et filiation), ce qui complique certaines requêtes (en particulier celles relatives à la relation spatio-temporelle).

La recherche proposée a permis de développer les premiers fondements nécessaires à l'intégration de la granularité spatiale dans les modélisations spatio-temporelles sous forme de graphe ou de bigraphe. Dans la première modélisation par un graphe spatio-temporel, nous avons montré qu'il était possible d'ajouter cette composante sous la forme d'une fonction (i.e. amalgamation) spécifiant la hiérarchie des différents niveaux de granularité, et que l'extraction des différents niveaux de détail se résume à une extraction de sous graphes.

La piste des bigraphes nous ouvre d'autres perspectives sur la façon de modéliser et de représenter explicitement un phénomène spatio-temporel selon plusieurs niveaux de granularité. Cette représentation permet d'identifier des blocs (i.e. site) que l'on peut éventuellement comparer si l'on recherche des similarités ou des différences entre deux phénomènes. Le bigraphe permet de conserver les liens de filiations entre entités même si seules certaines entités sont explicitement représentées (i.e. liens extérieurs). De plus, nous avons montré qu'il était possible de modéliser les changements par l'ajout de règles de dynamique au bigraphe. En comparaison à l'analyse des liens spatio-temporels d'un graphe, l'utilisation de ces règles pourrait améliorer l'efficacité du traitement informatique (e.g. par un analyseur syntaxique) pour identifier certains processus. On peut aussi imaginer traduire ces règles en phrases et reconstituer ainsi l'histoire d'un phénomène sous la forme de texte. Par contre il devient plus difficile d'établir certaines relations comme la relation spatio-temporelle (car le lien spatio-temporel n'existe plus explicitement), sinon par comparaison des différentes règles définies avant et après un changement dans la configuration spatiale (e.g. quelle entité est passée par tel endroit ?).

Conclusion

Cette étude de la problématique de représentation, de gestion et de manipulation de données spatio-temporelles a conduit au choix d'un modèle suffisamment souple et générique pour être appliqué à un grand nombre d'applications potentielles. Notre recherche se trouve à la frontière de plusieurs domaines liés en particulier aux sciences de l'information : la théorie des graphes, le raisonnement spatial et temporel et les Systèmes d'Information Géographique. Nous décrivons et réalisons la synthèse de notre contribution, puis dans un second temps nous présentons plusieurs perspectives qu'il reste à aborder afin de compléter et de poursuivre le travail de cette thèse.

Contribution

Partant du constat que :

- (1) l'espace et le temps sont indissociables et nécessaires à l'analyse de l'évolution d'entités spatio-temporelles,
- (2) les changements qui s'opèrent sur ces entités sont en particuliers caractérisés par des modifications spatiales (e.g. liées à la position, aux dimensions de l'empreinte spatiale) ou liées à leur identité,
- (3) les différents modèles décrits dans l'état de l'art ne gèrent pas de façon simultanée et complète, les relations relatives aux dimensions spatiale et temporelle ;

Nous proposons un *modèle de graphe* qui distingue les relations liées à l'espace et à l'identité, et ce, d'une part dans le temps (relations spatio-temporelles, filiations temporelles) et d'autre part à un instant donné (relations spatiales, filiations). Ce modèle est associé à des fonctions de manipulation qui permettent :

- d'extraire de l'information sur les entités du graphe par rapport à ces trois catégories de relations (e.g. fonctions de voisinage),

- d'identifier des processus particuliers (e.g. fonctions de vie et de mort) ou liés à une application,
- de combiner plusieurs sources d'informations (e.g. jointure selon les trois catégories de relations),
- de caractériser les relations entre entités par l'identification des différents types de chemins (i.e. route) du graphe qui peuvent exister entre ces entités (e.g. AllAll, AllSome). Ces fonctions permettent aussi d'identifier des processus de division (SomeAll) et de fusion (AllSome) entre deux temps consécutifs. Par ailleurs, étiqueter les routes de filiation par le résultat de la composition des relations appartenant à cette route, permet d'établir le type de la relation de filiation qui existe entre les entités présentes aux extrémités de cette route.

Nous avons ensuite abordé un volet plus technique en décrivant un *modèle de données* adapté au modèle de graphe spatio-temporel. La description de ce modèle de données nous permet de lever toute ambiguïté sur les constituants du modèle de graphe. En effet, il nous faut préciser notre précédente formalisation pour disposer de tous les paramètres nécessaires lors de la spécification de contraintes. Ces contraintes permettent de s'assurer : (1) de la cohérence (i.e. consistance) des données vis à vis du modèle de graphe avec la spécification de contraintes de *domaine*, et (2) de la cohérence des données par rapport à une application en particulier avec des contraintes dites *sémantiques*. Les quelques exemples de contraintes sémantiques que nous avons fournis ouvrent la voie à la modélisation d'autres contraintes, sur le même schéma, pour d'autres applications. De plus, afin de permettre une vérification de la consistance des données de façon automatique, nous avons proposé un algorithme de test de la consistance de ces contraintes sur une instance de graphe spatio-temporel.

Le dernier point abordé dans notre recherche part du constat que la représentation informatique d'un phénomène entraîne une description plus ou moins simplifiée de la réalité. Ceci implique que la modélisation et l'analyse sont fonction du niveau de détail disponible et/ou souhaité de cette description. Ainsi, dans l'objectif d'améliorer l'analyse d'un phénomène spatio-temporel décrit par notre modèle de graphe, nous avons entamé une réflexion sur l'intégration du concept de granularité spatiale (variation du niveau de détail) au sein de ce modèle. Nous avons montré comment exprimer la hiérarchie des différents niveaux de granularité spatiale par l'extraction de sous-graphes (e.g. fonction d'amalgamation). Nous avons ensuite eu l'opportunité d'étendre cette réflexion à l'utilisation d'une structure basée sur les bigraphes. Cette modélisation ouvre d'autres perspectives sur la manière de modéliser et de représenter explicitement un phénomène spatio-temporel selon plusieurs niveaux de granularité. Cette structure basée sur les bigraphes répond à des exigences spatio-temporelles proches de celles de notre modèle, mais où l'intégration de la granularité spatiale est optimisée :

- la granularité spatiale et les relations entre les différents niveaux de détails sont explicitement représentés,
- les liens de filiations entre entités sont conservés même si seules certaines entités sont explicitement représentées (i.e. liens extérieurs),
- il est possible de modéliser les changements par l'ajout de règles de dynamique au bigraphe.

Cependant, l'identification de certaines relations comme la relation spatio-temporelle est moins explicite, sinon par comparaison des différentes règles (e.g. quelle entité est passée par cet endroit ?).

Notre contribution abouti à la création d'un modèle de graphe spatio-temporel permettant de prendre en compte des relations spatiales et de filiations entre entités, avec une ouverture vers l'intégration de la granularité spatiale. A cette contribution est associé la spécification formelle d'un modèle de données et de contraintes. Un ensemble de fonctions est disponible pour l'étude de phénomènes du monde réel, l'adaptation à d'autres application est possible par l'ajout des contraintes et des fonctions nécessaires afin de répondre à des besoins spécifiques.

Perspectives

Approfondissement du concept de granularité spatiale

Outre l'intégration du concept de granularité spatiale au sein de notre modèle de graphe, l'introduction des bigraphes comme moyen d'explicitier les différents niveaux de granularité spatiale mérite d'être approfondie.

En particulier, il reste :

- à préciser son champ d'application et ses limites d'intérêt,
- à étendre la définition des bigraphes à toutes les relations du *RCC8*,
- à exploiter l'identification de sites (on peut éventuellement comparer ces "blocs" pour rechercher des similarités ou des différences entre deux phénomènes),
- à poursuivre la réflexion sur les règles de dynamiques : y a-t-il un intérêt majeur à utiliser de telles règles ? (e.g. efficacité de traitement) Si oui dans quelles perspectives ? (e.g. traduire le déroulement d'un phénomène sous la forme de texte, comparaison de la dynamique de plusieurs phénomènes, explicitation de processus spécifiques sous la forme de règles)

Approfondissement du concept de filiation

Nous avons décrit dans cette thèse deux principaux types de filiations : la continuation et la dérivation. Si la continuation est définie précisément comme une relation entre deux versions d'une même entité, la dérivation est laissée à l'appréciation de l'utilisateur du modèle. Il existe presque autant de dérivation que d'applications, il n'est donc pas question ici de les lister, mais plutôt d'établir des catégories. En particulier, il est utile de pouvoir distinguer le cas d'une absorption d'une fusion de deux éléments. Concrètement, faire la différence entre une personne qui mange un aliment, de deux parcelles que l'on fusionne.

Dans ce but, il s'agit de :

- spécifier formellement les grandes catégories de dérivation possibles,
- étudier les compositions entre ces différentes dérivation¹²,
- adapter le treillis relatif à la caractérisation des routes avec ces nouveaux types de dérivation,
- spécifier les contraintes émergentes à la suite de ces ajouts de types de relation au modèle.

Par ailleurs, il doit être possible de définir un *degré de filiation* permettant de qualifier le niveau de la relation entre deux entités. Par exemple, un frère et une sœur auront un degré plus élevé que celui qui existe entre deux cousins.

Complétion de la démarche spatio-temporelle globale du point de vue opérationnel

Nous avons abordé l'aspect opérationnel du traitement d'une problématique spatio-temporelle par la réalisation d'un prototype (Annexe I) reprenant les éléments du graphe spatio-temporel. Afin que ce travail soit complet, il reste les tâches qui suivent à terminer :

Contraintes et fonctions : Il s'agit de réaliser quelques aménagements entre la partie formelle sur le modèle de données et le prototype, ce dernier n'étant pas basé sur le même modèle de données. En particulier, lors de la spécification des contraintes, les relations spatiales ne sont pas stockées alors qu'elles le sont explicitement dans le modèle de données sur lequel est construit le développement.

Par ailleurs, afin de permettre une adaptation aisée à tout type d'application, il serait opportun d'intégrer au prototype une interface conviviale permettant l'ajout de fonctions, de contraintes, et

12. Nous l'avons fait pour dérivation/continuation avec par exemple la composition d'une dérivation avec une continuation nous décrivant globalement une dérivation.

de relations. Par exemple, on peut vouloir modifier l'ensemble des relations spatiales (actuellement RCC8), en ajoutant des relations du type loin/proche.

Concernant les fonctions génériques que propose le modèle, il faudrait ajouter les différentes fonctions de jointure qui ne sont pas disponibles actuellement dans le prototype.

Intégration de la granularité spatiale : On peut intégrer l'aspect granularité spatiale dans le prototype au travers des quelques pistes qui ont été proposées (e.g. fonction d'amalgamation). Concrètement, il faut prévoir l'extraction et la représentation des sous graphes inhérents à chaque niveau de granularité, et définir les fonctions relatives aux requêtes entre plusieurs niveaux de granularité.

Connexion à un Système d'Information Géographique : Afin d'avoir une vision plus "cartographique" du phénomène modélisé, quelques essais ont été faits pour proposer dans le prototype, l'affichage des polygones relatifs aux entités à partir du graphe. Cependant, il serait intéressant de permettre un accès direct de notre prototype à un logiciel de SIG afin de bénéficier à la fois de la vision sous forme de graphe et des fonctions particulières proposées par notre modèle, et de la vision cartographique ainsi que des fonctionnalités inhérentes à un SIG.

Application du modèle à une problématique issue de données réelles

L'étude, conceptuelle et théorique, de la problématique spatio-temporelle dans le cadre de l'analyse d'un phénomène réel, a été traitée de manière approfondie (i.e. formalisation d'un modèle de graphe, spécification de contraintes sur un modèle de données). Cependant, les cas applicatifs qui ont illustré cette démarche ne reposent pas sur des données réelles, l'échantillon des données est très petit, et il n'y a pas de problématique clairement identifiée sur le phénomène modélisé.

Par conséquent l'intérêt opérationnel du modèle reste à valider, en particulier sur le concept de routes qui pour l'instant n'a pas été exploité pleinement dans les cas d'études que nous avons pu décrire.

L'étape qui suit donc ce travail de modélisation est le traitement d'une problématique spatio-temporelle reposant sur des données réelles et utilisant le prototype développé (Annexe I). D'ores et déjà deux projets sont en cours d'élaboration, concernant respectivement l'évolution d'un champ de dunes marines et l'étude de la dynamique spatiale de données cadastrales.

L'évolution de dunes marines : Mené dans le cadre d'un partenariat avec le Service Hydrographique et Océanographique de la Marine (SHOM), ce projet¹³ concerne le suivi des positions et des relations spatiales de dunes sous marines sur six années (2002-2009) [Mascret, 2010]. Les données de départ sont des polygones représentant l'empreinte spatiale de ces dunes ce qui impose qu'à un instant donné, il n'est pas possible que deux dunes se chevauchent. En conséquence, le modèle de graphe a été étendu, l'ensemble Γ original comprenant uniquement les relations du *RCC8* devient un nouvel ensemble $\Gamma = \{EC, TPP, NTPP, PO, EQ, TPP_i, NTPP_i, AH_R, AH_L, BE_R, BE_L\}$, avec les relations AH_R et AH_L respectivement *devant droit* et *devant gauche*, BE_R et BE_L respectivement les relations *derrière droit* et *derrière gauche*.

Il nous reste à préciser avec nos correspondants du SHOM les éventuelles fonctions et contraintes supplémentaires à intégrer au modèle, pour analyser au mieux ce phénomène.

Etude de la dynamique spatiale de données cadastrales : Cette application est le fruit d'une collaboration en cours dans le cadre du GDR de Modélisation et Dynamiques Spatiales (MODYS) et du projet *MODELESPACE*, interdisciplinaire (association de deux laboratoires d'histoire et d'archéologie, et d'un laboratoire de mathématique) soutenu par l'Agence Nationale de la Recherche [Modelespace, 2010].

L'objectif général de ce projet est de développer une procédure d'étude permettant d'analyser les dynamiques spatiales des peuplements à partir de sources fiscales de différentes natures : des documents cadastraux scripturaux ou sous forme de plans, des photographies aériennes. Les données spatiales de ces sources pouvant être géolocalisées ou non.

Nous intervenons à un stade relativement avancé du projet où les entités (i.e. polygones relatifs à des parcelles) ont été extraites, et où se pose le choix du modèle spatio-temporel. Notre contribution consiste à intégrer l'ensemble de données cadastrales géolocalisées. Si le résultat des premiers tests est concluant, nous pourrions alors, en collaboration avec les membres du projet *MODELESPACE*, préciser les contraintes et fonctions nécessaires pour optimiser la modélisation et le traitement de la problématique.

13. Le début de ce travail est présenté en Annexe II.

Au terme de cette recherche, nous proposons un nouveau modèle spatio-temporel abordant l'apport de structures basées sur la théorie des graphes pour la représentation de la dynamique d'entités géographiques. L'approche permet au sein d'un même modèle de représenter à la fois des relations spatiales et des relations relatives à l'identité des entités [Del Mondo *et al.*, 2010]. De par l'aspect générique de notre formalisation, nous ouvrons potentiellement l'utilisation de ce modèle à un grand nombre d'applications. En effet, nous avons en particulier élaboré des schémas de fonctions et de contraintes nécessaires à l'adaptation de ce modèle à une problématique réelle [Del Mondo *et al.*, 2011]. D'un point de vue opérationnel, une première implémentation fonctionnelle du modèle a été réalisée. Il reste à valider le modèle sur des applications précises et améliorer notre prototype pour qu'il puisse être adapté facilement à ces applications. Le dernier chapitre de cette thèse ouvre la voie vers une réflexion autour du concept de granularité spatiale [Stell *et al.*, 2011], aspect encore peu exploré des théories de l'information spatiale.

Annexes

Annexe I : Implémentation du graphe spatio-temporel

Le modèle de graphe spatio-temporel présenté au chapitre II a donné lieu à une implémentation sous la forme d'une application en langage Java. Elle propose la visualisation du graphe spatio-temporel à partir d'une base de données et la plupart des fonctions de bases présentées dans le chapitre II (e.g voisinage, routes) sont disponibles.

Nous n'avons pas suivi le schéma destiné à formaliser les contraintes sur le modèle (Chapitre III) et toutes nos relations sont explicitement stockées dans des tables. Ceci car l'implémentation a eu lieu en amont de la formalisation du modèle de données pour l'établissement des contraintes. Le choix d'implémentation pour lequel nous avons opté n'est pas forcément moins bon que celui du modèle de données, en particulier il évite à chaque interrogation relative aux relations sur les entités de re-calculer ces relations ; elles sont déjà présentes dans la base.

La description complète de la base de données utilisée peut être consultée dans le rapport de [Lecointre, 2010]. Le détail de l'implémentation proprement dite est disponible dans le rapport de [Carpentier, 2010]. Nous présentons ici une vue générale de la base de données et la liste des fonctionnalités implémentées dans le prototype.

Description de la base de données

On dispose de quatre tables :

1. La table des **entités**- attributs :

- Id : identifiant unique de l'entité
- Géometrie : type d'objet géométrique géoréférencé
- Temps : position sur l'échelle des temps définie par le modèle
- Date : date d'acquisition des données

2. La table des **relations spatiales**- attributs :

- id_1 : identifiant de la première entité liée par la relation spatiale
- id_2 : identifiant de la deuxième entité liée par la relation spatiale
- relation : indique le type de relation qui lie les deux entités (e.g *RCC8*)

3. La table des **relations de filiation**- attributs :

- id_1 : identifiant de la première entité liée par la relation de filiation
 - id_2 : identifiant de la deuxième entité liée par la relation de filiation
 - relation : indique le type de relation qui lie les deux entités (e.g continuation, dérivation)
4. La table des **relations spatio-temporelles**- attributs :
- id_1 : identifiant de la première entité liée par la relation spatio-temporelle
 - id_2 : identifiant de la deuxième entité liée par la relation spatio-temporelle
 - relation : indique le type de relation qui lie les deux entités (e.g *RCC8*)

Fonctionnalités implémentées

Le prototype dispose des fonctionnalités suivantes :

↔ Représentation

- Afficher un graphe spatio-temporel
- Personnaliser l'affichage (e.g couleurs)
- Editer le graphe (e.g suppression d'entités/arcs) avec possibilité d'enregistrer les changements dans la base de données

↔ Analyse

- Fonction de voisinage spatial avec choix de la distance
- Fonction ascendance et descendance pour les relations de filiation et spatio-temporelle
- Fonction qui teste si deux entités sont en relation spatio-temporelle entre deux temps t et $t + n$, permet de vérifier la transitivité de la relation spatio-temporelle.
- Fonction qui qualifie (e.g All/All,Some/All) une route entre deux entités
- Fonction qui renvoie les potentielles entités en commun entre deux routes
- Fonction qui à partir d'une entité e et d'un type de route, renvoie les entités qui entretiennent ce type de route avec e
- Calcul de la proportion de chacun des types de route dans le graphe entier

Annexe II : Application aux données de dunes marines

Nous présentons dans cette annexe un travail en cours de collaboration avec le SHOM. L'étude de la dynamique des dunes sous-marines résulte entre autres de besoins en géosciences marines et relatifs à la sécurité de la navigation.

Les données de dunes marines que nous avons utilisées ont été acquises dans la zone définie Figure 83. La Figure 84 représente les données bathymétriques d'un champ de dunes marines de cette zone entre les années 2002 et 2009. L'extraction des paramètres attributaires de ces dunes est effectuée à l'aide du logiciel ParamDune (Figure 85). Les 30 dunes identifiées sur les 6 enregistrements (de 2002 à 2009) sont visualisées sous ArcGis (Figure 86), et un exemple d'intégration de ces données dans notre prototype est proposé Figure 87. Le graphe spatio-temporel ainsi défini, et une étroite coopération avec les experts du SHOM doit nous aider à caractériser la dynamique de ces dunes marines. D'un point de vue plus général, l'objectif est de pouvoir étudier des champs de dunes plus complexes, comme par exemple celui du *Banc du four* localisé aux environs de Ouessant (Figure 88).

A ce stade de la collaboration avec le SHOM, nous avons récupéré un petit échantillon de données de trente dunes que nous avons intégré dans notre prototype. Dans un premier temps, ceci nous permet de valider l'intégration de données réelles dans ce prototype, ainsi que l'implémentation de ses différentes fonctionnalités¹⁴. L'utilisation de ces opérateurs pour l'analyse de processus n'a pas été intégrée à ce travail, elle devra l'être lorsque nous traiterons de l'ensemble de données du *Banc du four* plus complexe d'environ 500 dunes.

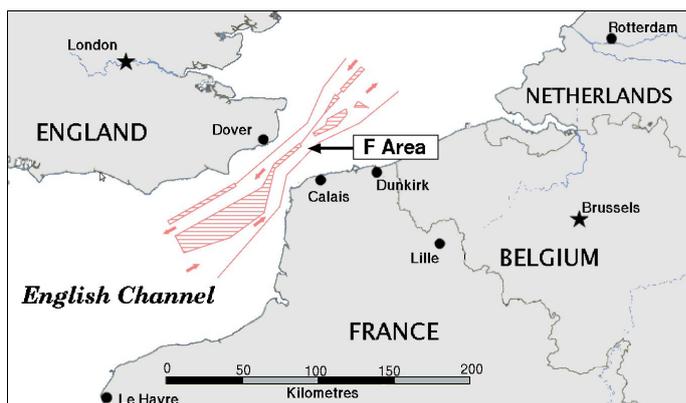


Figure 83 – Zone d'étude

14. Fonctionnalités présentées dans l'annexe I.

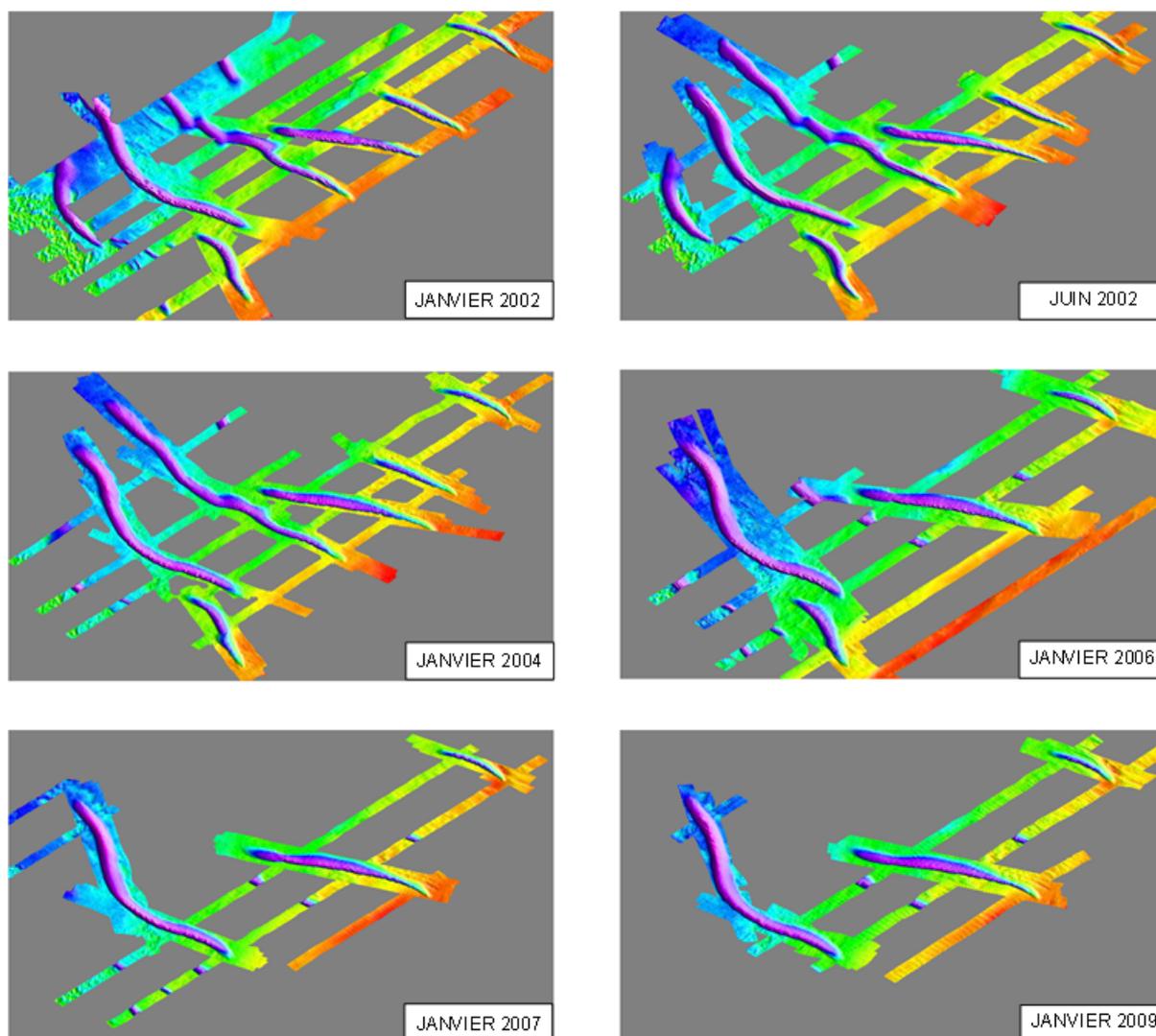


Figure 84 – Données bathymétriques

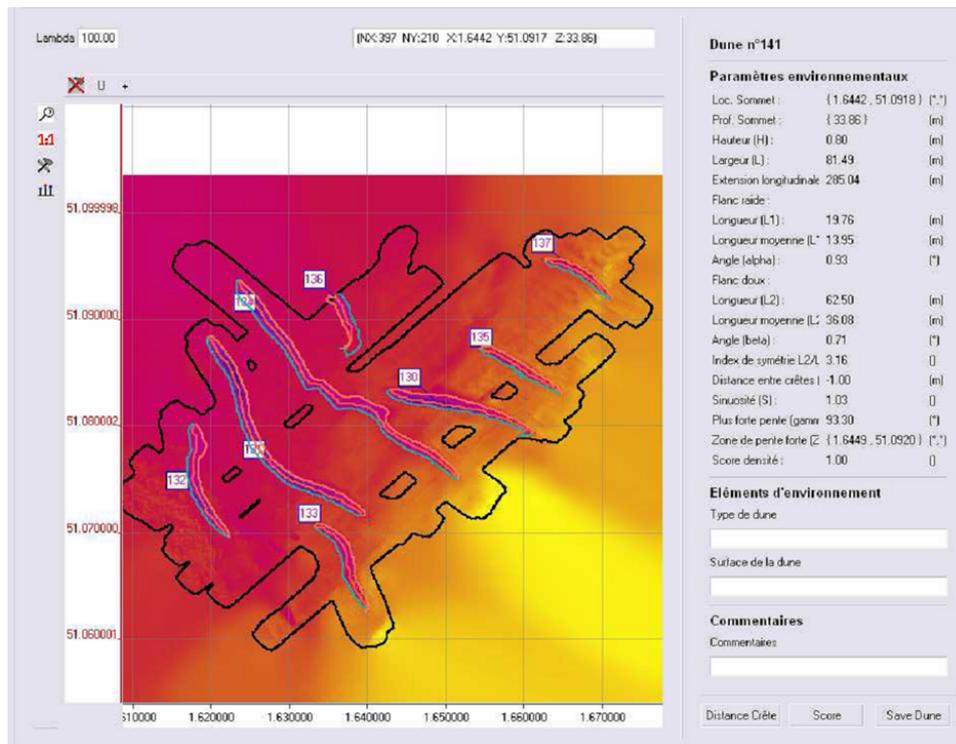


Figure 85 – Extraction automatique des paramètres attributaires à l'aide du logiciel ParamDune.



Figure 86 – Représentation des données dans le logiciel ArcGis

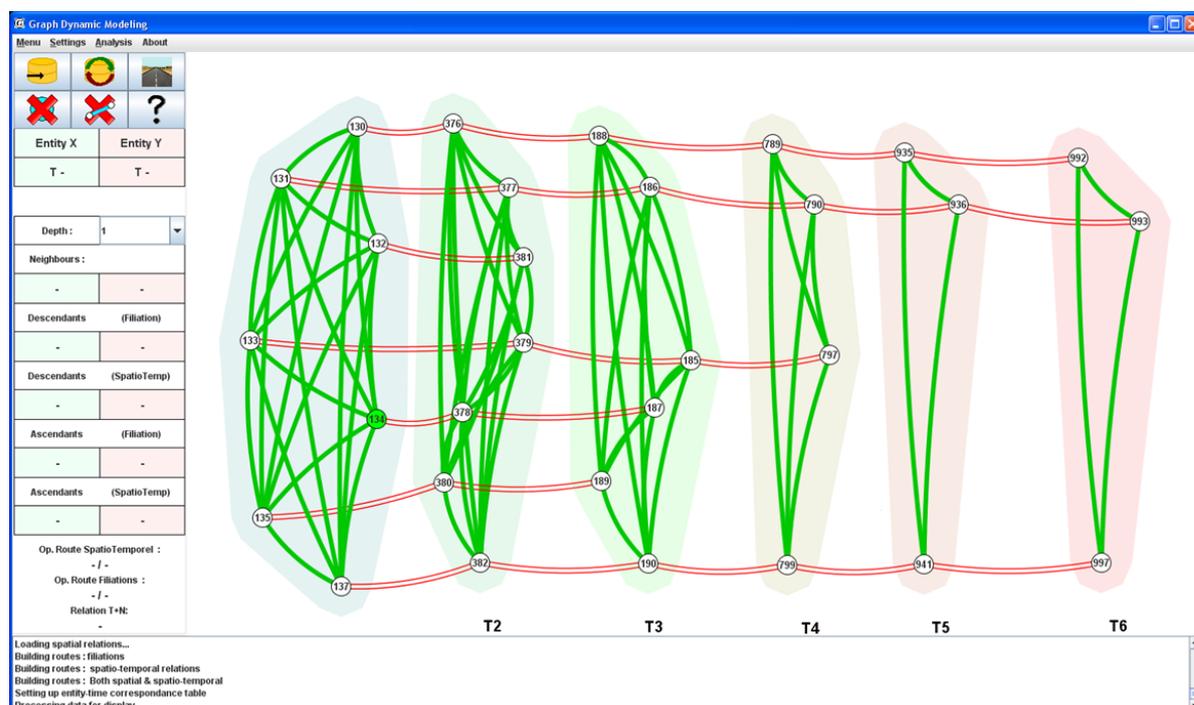
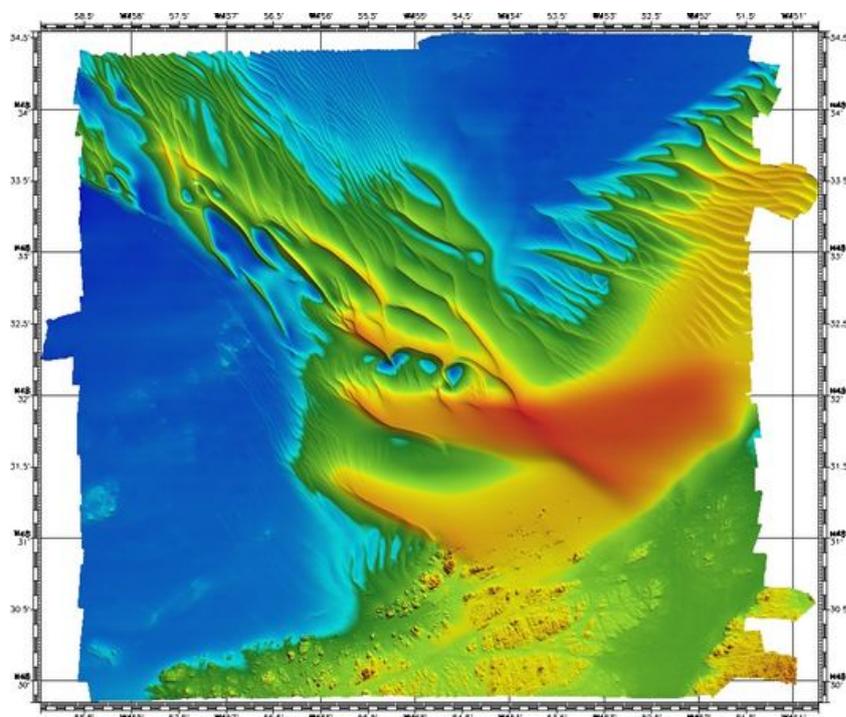


Figure 87 – Intégration des données de dunes dans le prototype

Figure 88 – Un exemple de champ de dunes plus complexe (*Banc du four*) - Image IFREMER

Annexe III : Preuve de consistance d'un ensemble de contraintes

Cette preuve est extraite de l'article [Del Mondo *et al.*, 2011]. Les contraintes citées dans la preuve (respectivement de type RD et FC) sont rappelées ci-dessous. Dans la thèse elles correspondent respectivement aux équations III.2 et III.3. Par ailleurs les contraintes EIDs, TFD et GFDs évoquées au début de la preuve sont relatives aux équations III.4, III.5, III.6 et au paragraphe III.3.2.

Let $G_{sp} = \langle V, E \rangle$ be a spatio-temporal graph schema with $O, O' \in V$ and $R \in E$.

(RD) A filiation predicates holds an inclusion dependency with respect to spatio-temporal predicates :

$$\begin{aligned} & \forall id_1 id_2 t_1 t_2 \rho (R(id_1, t_1, \rho, id_2, t_2)) \\ & \rightarrow \exists \bar{u}_1, g_1, \bar{u}_2, g_2 (O(id_1, t_1, \bar{u}_1, g_1) \wedge O'(id_2, t_2, \bar{u}_2, g_2)). \end{aligned} \quad (1)$$

(FC) Then, the following filiation constraint must hold :

$$\bar{\forall} (R(id_1, t_1, \rho, id_2, t_2) \rightarrow (id_1 \neq id_2) \wedge (t_1 \leq t_2)). \quad (2)$$

Satisfiability

Satisfiability is a classical reasoning problem that applies to integrity constraints. Following the notion of satisfiability in databases, a set of integrity constraints ψ will be satisfiable if there exists a non-empty instance D of a spatio-temporal graph schema G that satisfies ψ , which is denoted by $D \models \psi$.

proposition Given a spatio-temporal graph schema $G_{sp} = \langle V, E \rangle$ of Σ with domain constraints RD and FC and a set ψ of EIDs, TFD and GFDs, then ψ is always satisfiable.

proof We prove this proposition by showing that there always exists an instance D with a single tuple in an object predicate such that $D \models \psi$.

RDs impose inclusion dependencies (Equation 1), which are acyclic; that is, the consequence of the inclusion dependency cannot be the antecedent in another inclusion dependency. These are the only existential constraints (i.e., they are defined by an existential quantification) in ψ . So, it is enough to include only a tuple in an object predicate so that the instance is non-empty and the inclusion dependencies are satisfied. A database instance with only a tuple in an object predicate is equivalent to a graph with a single node.

By the filiation constraints defined in equation 2, a derivation needs different identities. So, no single tuple of an object predicate can participate in a filiation relation, and consequently, neither FCs nor EIDs, TFDs and GFDs can be triggered. Therefore, with a single tuple in an object predicate we have a non-empty database instance that is consistent. \square

Notice that this notion of satisfiability for the set of ψ does not guarantee that a set of EIDs, TFDs and GFDs is free of contradictions. For example, the following TFD, with a constant value $\rho_{f'}$ contradict each other.

Bibliographie

- [Allen, 1984] ALLEN, J. F. (1984). Towards a general theory of action and time. *Artif. Intell.*, 23(2) :123–154.
- [Allen et Ferguson, 1994] ALLEN, J. F. et FERGUSON, G. (1994). Actions and events in interval temporal logic. *Journal of Logic and Computation*, 4(5) :531–579.
- [Allen et Hayes, 1985] ALLEN, J. F. et HAYES, P. J. (1985). A common-sense theory of time. In *IJCAI*, pages 528–531.
- [Allen et Hayes, 1990] ALLEN, J. F. et HAYES, P. J. (1990). Moments and points in an interval-based temporal logic. *Comput. Intell.*, 5 :225–238.
- [Armstrong, 1988] ARMSTRONG, M. (1988). Temporality in spatial databases. In *Proceedings of GIS/LIS'88*, volume 2, pages 880–889, Bethesda, Maryland.
- [Barrera et al., 1991] BARRERA, R., FRANK, A. et AL-TAHA, K. (1991). Temporal relations in geographic information systems : A workshop at the university of maine, orono, october 12-13, 1990. *ACM Sigmod Record*, 20 :85–91.
- [Batty, 2003] BATTY, M. (2003). Network geography : Relations, interactions, scaling and spatial processes in gis. *UCL Working Papers Series*, 63. ISSN 1467-1298.
- [Batty, 2004] BATTY, M. (2004). A new theory of space syntax. CASA Working Papers 75, Centre for Advanced Spatial Analysis, University College London.
- [Bejaoui et al., 2007] BEJAOU, L., BÉDARD, Y., PINET, F. et SCHNEIDER, M. (2007). Logical consistency for vague spatiotemporal objects and relations. In *The 5th International Symposium on Spatial Data Quality (ISSDQ 2007)*.
- [Bestougeff et Ligozat, 1992] BESTOUGEFF, H. et LIGOZAT, G. (1992). *Logical tools for temporal knowledge representation*. Ellis Horwood, UK.

- [Billen et Clementini, 2006] BILLEN, R. et CLEMENTINI, E. (2006). Projective relations in a 3d environment. In RAUBAL, M., MILLER, H., FRANK, A. et GOODCHILD, M., éditeurs : *Geographic Information Science*, volume 4197 de *Lecture Notes in Computer Science*, pages 18–32. Springer Berlin Heidelberg.
- [Bordin, 2010] BORDIN, P. (2010). L'intégration du temps dans les sig. *XYZ*, (122).
- [Brisaboa et al., 1998] BRISABOA, N., MIRBEL, I. et PERNICE, B. (1998). Constraints in spatio-temporal databases : A proposal of classification. In *Workshop on Exploring Modeling Methods in Systems Analysis and Design*.
- [Camossi, 2004] CAMOSSO, E. (2004). *Spatio-temporal multigranularity in an object data model*. Thèse de doctorat, Université de Milan.
- [Carpentier, 2010] CARPENTIER, C. (2010). Implémentation d'un modèle destiné à l'étude de l'évolution d'entités géographiques au cours du temps. Rapport technique, Institut de Recherche de l'École Navale et l'Institut Supérieur d'Informatique, de Modélisation et de leurs Applications.
- [Claramunt et Jiang, 2000] CLARAMUNT, C. et JIANG, B. (2000). Hierarchical reasoning in time and space. In *Proceedings 9th International Symposium on Spatial Data Handling*. P. Forer, A. G. O. Yeh and J. He (eds.).
- [Claramunt et Parent, 2003] CLARAMUNT, C. et PARENT, C. (2003). Modelling concepts for the representation of evolution constraints. *Computers, Environment and Urban Systems*, 27(3) :225–241.
- [Claramunt et Thériault, 1995] CLARAMUNT, C. et THÉRIAULT, M. (1995). Managing time in GIS : An event-oriented approach. In CLIFFORD, J. et TUZHILIN, A., éditeurs : *Proceedings of the VLDB International Workshop on Temporal Databases*, pages 23–42, Zurich, Switzerland. Workshops in Computing Series, Springer Verlag.
- [Claramunt et Thériault, 1996] CLARAMUNT, C. et THÉRIAULT, M. (1996). Toward semantics for modelling spatio-temporal processes within gis. In *Advances in GIS II*, pages 47–64, Delft, the Netherlands. Taylor and Francis.
- [Clifford et Croker, 1988] CLIFFORD, J. et CROKER, A. (1988). Objects in time. *Information Systems Working Papers Series*.
- [Cohn et al., 1997] COHN, A. G., BENNETT, B., GOODAY, J. et GOTTS, N. M. (1997). Qualitative spatial representation and reasoning with the region connection calculus. *Geoinformatica*, 1(3) :275–316.

- [Cohn et Renz, 2007] COHN, A. G. et RENZ, J. (2007). Qualitative spatial representation and reasoning. In van HARMELEN, F., LIFSCHITZ, V. et PORTER, B., éditeurs : *Handbook of Knowledge Representation*. Elsevier.
- [Cohn et Varzi, 1999] COHN, A. G. et VARZI, A. C. (1999). Modes of connection. In *Spatial Information Theory. Proceedings of the Fourth International Conference, Berlin and*, pages 299–314. Springer-Verlag.
- [Couclelis, 1992] COUCLELIS, H. (1992). People manipulate objects (but cultivate fields) : Beyond the raster-vector debate in gis. In FRANK, A. U., CAMPARI, I. et FORMENTINI, U., éditeurs : *Spatio-Temporal Reasoning - Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, International Conference GIS - From Space to Territory : Theories and Methods of Spatio-Temporal Reasoning, Pisa, Italy, September 21-23, 1992, Proceedings*, volume 639 de *Lecture Notes in Computer Science*, pages 65–77. Springer.
- [Couclelis et Liu, 2000] COUCLELIS, H. et LIU, X. (2000). The geography of time and ignorance : dynamics and uncertainty in integrated urban-environmental process models. In *Proceedings of the 4th International Conference on Integrating Geographic Information Systems and Environmental Modeling : Problems, Prospects and Needs for Research*, Alberta, Canada. K. C. Clarke and M. P. Crane (eds.).
- [Del Mondo *et al.*, 2011] DEL MONDO, G., RODRIGUEZ, M. A., CLARAMUNT, C. et THIBAUD, T. (2011). Modelling consistency of spatio-temporal graphs. *Soumis à : Data and Knowledge Engineering*.
- [Del Mondo *et al.*, 2010] DEL MONDO, G., STELL, J. G., CLARAMUNT, C. et THIBAUD, R. (2010). A graph model for spatio-temporal evolution. *Journal of Universal Computer Science*, 16(11) :1452–1477.
- [Duckham *et al.*, 2010] DUCKHAM, M., STELL, J. G., VASARDANI, M. et WORBOYS, M. F. (2010). Qualitative change to 3-valued regions. In *GIScience*, pages 249–263.
- [Eckmann et Boulton, 2008] ECKMANN, M. et BOULT, T. E. (2008). Spatio-temporal consistency and distributivity as qualities of features. In *Computer Vision and Pattern Recognition Workshops*.
- [Egenhofer, 1989] EGENHOFER, M. J. (1989). A formal definition of binary topological relationships. In *3rd International Conference, FODO 1989 on Foundations of Data Organization and Algorithms*, pages 457–472, New York, NY, USA. Springer-Verlag New York, Inc.
- [Egenhofer et Al-Taha, 1992] EGENHOFER, M. J. et AL-TAHA, K. K. (1992). Reasoning about gradual changes of topological relationships. In *Proceedings of the International Conference GIS*

- *From Space to Territory : Theories and Methods of Spatio-Temporal Reasoning on Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, pages 196–219, London, UK. Springer-Verlag.
- [Frank, 1994] FRANK, A. U. (1994). Qualitative temporal reasoning in gis - ordered time scales. In WAUGH, T. C. et (EDS.), R. C. H., éditeurs : *Proceedings of the Sixth International Symposium on Spatial Data Handling*, pages 410–430. Taylor & Francis.
- [Frank, 1996] FRANK, A. U. (1996). Qualitative spatial reasoning : Cardinal directions as an example. *International Journal of Geographical Information Science*, 10(3) :269–290.
- [Frank, 2001] FRANK, A. U. (2001). *Life and Motion of Socio-Economic Units*. A U Frank, J Raper et J P Cheylan.
- [Frank et Timpf, 1994] FRANK, A. U. et TIMPF, S. (1994). Multiple representations for cartographic objects in a multi-scale tree - an intelligent graphical zoom. *Computers & Graphics*, 18(6) :823–829.
- [Freksa, 1992a] FREKSA, C. (1992a). Temporal reasoning based on semi-intervals. *Artif. Intell.*, 54(1) :199–227.
- [Freksa, 1992b] FREKSA, C. (1992b). Using orientation information for qualitative spatial reasoning. In *Spatio-Temporal Reasoning*, pages 162–178.
- [Galton, 2000] GALTON, A. (2000). *Qualitative Spatial Change*. Oxford University Press.
- [Galton, 2004] GALTON, A. (2004). Fields and objects in space, time and space-time. *Spatial Cognition and Computation*, 4(1) :1–end.
- [Goodchild, 1989] GOODCHILD, M. F. (1989). Tiling large geographical databases. In *SSD'89*, pages 137–146.
- [Goodchild, 2010] GOODCHILD, M. F. (2010). A geospatial perspective. Position paper in Spatio-Temporal Constraints on Social Networks Specialist Meeting.
- [Goodchild et al., 1992] GOODCHILD, M. F., HAINING, R. et WISE, S. (1992). Integrating gis and spatial data analysis : problems and possibilities. *International Journal of Geographical Information Systems*, 6 :407–423.
- [Haddad, 2009] HADDAD, H. (2009). *Une approche pour supporter l'analyse qualitative des suites d'actions dans un environnement géographique virtuel et dynamique*. Thèse de doctorat, Université Laval.
- [Hägerstrand, 1970] HÄGERSTRAND, T. (1970). What about people in regional science. *Papers of the Regional Science Association*, 24 :6–21.

- [Hallot et Billen, 2008] HALLOT, P. et BILLEN, R. (2008). Life and motion configurations : A basis for spatio-temporal generalized reasoning model. *In ER Workshops*, pages 323–333.
- [Hillier et Hanson, 1984] HILLIER, B. et HANSON, J. (1984). *The Social Logic of Space*. Cambridge University Press, Cambridge, reprint édition.
- [Hirsch et Hodkinson, 2002] HIRSCH, R. et HODKINSON, I. (2002). *Relation Algebras by Games*. North-Holland.
- [Hornsby et Egenhofer, 1997] HORNSBY, K. et EGENHOFER, M. J. (1997). Qualitative representation of change. *In Spatial Information Theory, LNCS*, volume 1257, pages 15–33. Springer-Verlag.
- [Hornsby et Egenhofer, 2000] HORNSBY, K. et EGENHOFER, M. J. (2000). Identity-based change : a foundation for spatio-temporal knowledge representation. *International Journal of Geographical Information Science*, 14(3) :207–224.
- [Hornsby et Egenhofer, 2002] HORNSBY, K. et EGENHOFER, M. J. (2002). Modeling moving objects over multiple granularities. *Annals of Mathematics and Artificial Intelligence*, 36(1-2) :177–194.
- [Jensen *et al.*, 1994] JENSEN, C. S., CLIFFORD, J., ELMASRI, R., GADIA, S. K., HAYES, P. J. et JAJODIA, S. (1994). A consensus glossary of temporal database concepts. *SIGMOD Record*, 23(1) :52–64.
- [Jiang et Claramunt, 2004a] JIANG, B. et CLARAMUNT, C. (2004a). A structural approach to the model generalization of an urban street network. *GeoInformatica*, 8(2) :157–171.
- [Jiang et Claramunt, 2004b] JIANG, B. et CLARAMUNT, C. (2004b). Topological analysis of urban street networks. *Environment and Planning B : Planning and Design*, 31(1) :151–162.
- [Jiang *et al.*, 2000] JIANG, B., CLARAMUNT, C. et KLARQVIST, B. (2000). An integration of space syntax into gis for modelling urban spaces. *Journal of American Geographers*, 2 :161–171.
- [Jiang et Worboys, 2009] JIANG, J. et WORBOYS, M. (2009). Event-based topology for dynamic planar areal objects. *Int. J. Geogr. Inf. Sci.*, 23(1) :33–60.
- [Khoshafian et Copeland, 1986] KHOSHAFIAN, S. et COPELAND, G. P. (1986). Object identity. *In Object-Oriented Programming Systems, Languages, and Applications (OOPSLA '86)*, pages 406–416.
- [Kurata, 2009] KURATA, Y. (2009). Semi-automated derivation of conceptual neighborhood graphs of topological relations. *In COSIT*, pages 124–140.
- [Kurata et Egenhofer, 2006] KURATA, Y. et EGENHOFER, M. J. (2006). The head-body-tail intersection for spatial relations between directed line segments. *In GIScience Geographic Information*

- Science, 4th International Conference, GIScience 2006, Münster, Germany, September 20-23, 2006, Proceedings*, volume 4197 de *Lecture Notes in Computer Science*, pages 269–286. Springer.
- [Ladkin, 1987] LADKIN, P. B. (1987). *The Logic of Time Representation*. Thèse de doctorat, University of California at Berkeley.
- [Langran, 1992] LANGRAN, G. (1992). *Time in Geographic Information Systems*. Taylor & Francis.
- [Langran et Chrisman, 1988] LANGRAN, G. et CHRISMAN, N. (1988). A framework for temporal geographic information. *JournalCartographica : The International Journal for Geographic Information and Geovisualization*, 25(3) :1–14.
- [Lecointre, 2010] LECOINTRE, A. (2010). Chaîne de traitement et d'analyse pour la caractérisation de la dynamique des reliefs. Rapport technique, Institut de Recherche de l'École Navale et l'École Supérieure des Géomètres et Topographes.
- [Ligozat, 1998] LIGOZAT, G. (1998). Reasoning about cardinal directions. *Journal of Visual Languages and Computing*, (9) :315–324.
- [Mascret, 2010] MASCRET, A. (2010). Chaîne de traitement pour la caractérisation de la dynamique des dunes sous-marines. Rapport technique, Rapport d'étude - Service Hydrographique et Océanographique de la Marine.
- [Medak, 1999] MEDAK, D. (1999). Lifestyles - an algebraic approach to change in identity. In *STDBM '99 : Proceedings of the International Workshop on Spatio-Temporal Database Management*, pages 19–38, London, UK. Springer-Verlag.
- [Meentemeyer, 1989] MEENTEMEYER, V. (1989). Geographical perspectives of space, time, and scale. *Landscape Ecology*, 3(3) :163–173.
- [Milner, 2009] MILNER, R. (2009). *The Space and Motion of Communicating Agents*. Cambridge University Press.
- [Modelespace, 2010] MODELESPACE (2010). Projet de recherche modelespace 2010-2012. <http://w3.modele-espace.univ-tlse2.fr/>.
- [Muller, 2002] MULLER, P. (2002). Topological spatio-temporal reasoning and representation. *Computational Intelligence*, 18(3) :420–450.
- [Open GIS Consortium Inc., 1999] OPEN GIS CONSORTIUM INC. (1999). OpenGIS Simple Features Specification for SQL. Rapport technique, Open GIS Consortium.
- [Ott et Swiaczny, 2001] OTT, T. et SWIACZNY, F. (2001). *Time-Integrative Geographic Information Systems*. Springer.

- [Parkes et Thrift, 1980] PARKES, D. et THRIFT, N. (1980). *Times, spaces and places : A chrono-geographic perspective*.
- [Peuquet, 1984] PEUQUET, D. (1984). A conceptual framework and comparison of spatial data models. *Cartographica The International Journal for Geographic Information and Geovisualization*, 21(4) :66.
- [Peuquet, 2002] PEUQUET, D. (2002). *Representations of Space and Time*. Guilford Press, New York.
- [Peuquet et Duan, 1995] PEUQUET, D. et DUAN, N. (1995). An event-based spatiotemporal data model (estdm) for temporal analysis of geographical data. *International Journal of Geographical Information Systems*, 9(1) :7–24.
- [Peuquet et Wentz, 1994] PEUQUET, D. et WENTZ, E. (1994). An approach time-based analysis of spatiotemporal data. In UNION, I. G., éditeur : *Sixth International Symposium on Spatial Data Handling*, volume 1, pages 489–504, Edinburgh, Scotland.
- [Pullar et Egenhofer, 1988] PULLAR, D. V. et EGENHOFER, M. J. (1988). Towards the defaction and use of topological relations among spatial objects. In *Proceedings of the 3rd International Symposium on Spatial Data Handling. IGU, Colombus*, pages 225–242.
- [Randell et al., 1992] RANDELL, D. A., CUI, Z. et COHN, A. G. (1992). A spatial logic based on regions and connection. In NEBEL, B., RICH, C. et SWARTOUT, W., éditeurs : *KR'92. Principles of Knowledge Representation and Reasoning : Proceedings of the Third International Conference*, pages 165–176, San Mateo, California. Morgan Kaufmann.
- [Rodríguez et al., 2008] RODRÍGUEZ, M. A., BERTOSSI, L. et CANIUPÁN, M. (2008). An inconsistency tolerant approach to querying spatial databases. In *GIS '08 : Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*, pages 1–10, New York, NY, USA. ACM.
- [Rodríguez et al., 2010] RODRÍGUEZ, M. A., BRISABOA, N., MEZA, J. et LUACES, M. R. (2010). Measuring consistency with respect to topological dependency constraints. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '10*, pages 182–191, New York, NY, USA. ACM.
- [Saux, 2010] SAUX, E. (2010). Contribution à la représentation des espaces géographiques par des approches quantitatives et qualitatives. Habilitation à Diriger les Recherches, Université de Rennes 1.
- [Schlieder, 1995] SCHLIEDER, C. (1995). Reasoning about ordering. In FRANK, A. et KUHN, W. e., éditeurs : *COSIT*, volume 988 de *Lecture Notes in Computer Science*, pages 341–349. Springer.

- [Shu *et al.*, 2006] SHU, H., SPACCAPIETRA, C., PARENT, C. et QUESADA, S. (2006). Modelling uncertainty in spatiotemporal objects. *In Proceedings Of Accuracy 2006*, pages 469–478.
- [Smith, 1994] SMITH, B. (1994). Fiat objects. *Topoi Journal*, 20.
- [Snodgrass et Ahn, 1986] SNODGRASS, R. T. et AHN, I. (1986). Temporal databases. *IEEE Computer*, 19(9) :35–42.
- [Spéry *et al.*, 2001] SPÉRY, L., CLARAMUNT, C. et LIBOUREL, T. (2001). A spatio-temporal model for the manipulation of lineage metadata. *Geoinformatica*, 5(1) :51–70.
- [Sriti *et al.*, 2005] SRITI, M., THIBAUD, R. et CLARAMUNT, C. (2005). A fuzzy identity-based temporal gis for the analysis of geomorphometry changes. *Journal of Data Semantics*, 1(3) :81–99. Special issue on spatial data semantics.
- [Stefanakis, 2003] STEFANAKIS, E. (2003). Modelling the history of semi-structured geographical entities. *International Journal of Geographical Information Science*, 17(6) :517–546.
- [Stell, 1999] STELL, J. G. (1999). Granulation for graphs. *In Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. International Conference COSIT'99, volume 1661 of Lecture Notes in Computer Science*, pages 417–432. Springer-Verlag.
- [Stell, 2003] STELL, J. G. (2003). Granularity in change over time. *In DUCKHAM, M., GOODCHILD, M. et WORBOYS, M., éditeurs : Foundations of Geographic Information Science*, pages 95–115. Taylor and Francis.
- [Stell *et al.*, 2011] STELL, J. G., DEL MONDO, G., THIBAUD, R. et CLARAMUNT, C. (2011). Spatio-temporal evolution as bigraph dynamics. *In Conference on Spatial Information Theory (COSIT '11)*, pages 148–167. Springer LNCS 6899.
- [Stell et Worboys, 1999] STELL, J. G. et WORBOYS, M. F. (1999). Generalizing graphs using amalgamation and selection. *In Advances in Spatial Databases. 6th International Symposium, SSD'99, volume 1651 of Lecture Notes in Computer Science*, page pages. Springer.
- [Sundarraaj *et al.*, 2006] SUNDARRAJ, D., GIBBONS, P. et PILLAI, P. (2006). Ensuring spatio-temporal consistency in distributed networks of smart cameras. *In In Distributed Systems of Cameras Workshop at ACM Sensys*.
- [Thériault *et al.*, 1999] THÉRIAULT, M., CLARAMUNT, C. et VILLENEUVE, P. Y. (1999). A spatio-temporal taxonomy for the representation of spatial set behaviours. *In Spatio-Temporal Database Management*, pages 1–18.

- [Thériault et Claramunt, 1999] THÉRIAULT, M. et CLARAMUNT, C. (1999). La modélisation du temps et des processus dans les sig : Un moyen d'intégration pour la recherche interdisciplinaire. *Revue Internationale de Geomatique*, 9(1) :67–103.
- [Thériault et al., 2002] THÉRIAULT, M., CLARAMUNT, C., SEGUIN, A. M. et VILLENEUVE, P. (2002). Temporal gis and statistical modelling of personal lifelines. *In Proceedings of the 9th Spatial Data Handling Symposium*.
- [Vangenot, 2004] VANGENOT, C. (2004). Multi-representation in spatial databases using the mads conceptual model. *In ICA Workshop on Generalisation and Multiple representation*.
- [Wehrle, 1985] WEHRLER, L. (1985). La ronce (*rubus fruticosus* l. agg.) en forêt. *Revue Forestière Française*, 37(4) :288–304. ISSN 0035-2829 ENGREF, Ecole nationale du génie rural, des eaux et des forêts, Nancy (FRA).
- [Worboys, 2005] WORBOYS, M. (2005). Event-oriented approaches to geographic phenomena. *International Journal of Geographical Information Science*, 19(1) :1–28.
- [Worboys, 2010] WORBOYS, M. (2010). Responsive social networks. Position paper in Spatio-Temporal Constraints on Social Networks Specialist Meeting.
- [Worboys, 1990] WORBOYS, M. F. (1990). Reasoning about gis using temporal and dynamic logics. Rapport technique 90-4, NCGIA.
- [Worboys, 1992] WORBOYS, M. F. (1992). A model for spatio-temporal information. *In Proceedings 5th International Symposium on Spatial Data Handling*, volume 2, pages 602–611. IGU Commission of GIS.
- [Worboys, 1994] WORBOYS, M. F. (1994). A unified model for spatial and temporal information. *Comput. J.*, 37(1) :36–34.
- [Worboys, 2001] WORBOYS, M. F. (2001). Nearness relations in environmental space. *International Journal of Geographical Information Science*, 15.
- [Worboys et Hornsby, 2004] WORBOYS, M. F. et HORNSBY, K. (2004). From objects to events : Gem, the geospatial event model. *In GIScience*, pages 327–344.
- [Yuan, 1996] YUAN, M. (1996). Temporal gis and spatio-temporal modeling. *In Third International Conference/Workshop on Integrating GIS and Environmental Modeling*. Classeur 2.
- [Yuan, 1999] YUAN, M. (1999). Use of a three-domain representation to enhance gis support for complex spatiotemporal queries. *Transaction in GIS*, 3(2) :137–159.
- [Zimmermann et Freksa, 1996] ZIMMERMANN, K. et FREKSA, C. (1996). Qualitative spatial reasoning using orientation, distance, and path knowledge. *Applied Intelligence*, 6 :49–58.

UN MODELE DE GRAPHE SPATIO-TEMPOREL POUR REPRESENTER L'EVOLUTION D'ENTITES GEOGRAPHIQUES

Résumé

L'espace et le temps sont des concepts indissociables et nécessaires à l'analyse de l'évolution d'entités spatio-temporelles. Modéliser ces évolutions passe par la détection des changements et des relations qui modifient et caractérisent ces entités. Ces changements sont en particulier caractérisés par des modifications liées à la position, à l'empreinte spatiale de ces entités, ou même à leur identité.

Une modélisation spatio-temporelle nécessite le développement de solutions de représentation qui permettent d'identifier et d'étudier les processus mis en jeu, leurs causes et leurs conséquences. Le fait que la représentation formelle et informatique d'un phénomène entraîne une description plus ou moins simplifiée de la réalité, implique qu'une telle modélisation retienne le niveau de détail disponible et/ou souhaité de cette description. Idéalement, il devrait être possible d'étudier un phénomène à plusieurs niveaux de granularité dans l'espace et dans le temps. Si de nombreux modèles et approches ont été proposés pour modéliser des évolutions d'entités au sein de phénomènes spatio-temporels, aucun ne permet de disposer d'une structure de représentation permettant de saisir complètement la sémantique de ces applications.

Cette recherche propose un modèle de graphe spatio-temporel qui permet de caractériser les principales propriétés de l'évolution d'entités spatiales. Nous distinguons dans ce cadre plusieurs concepts structurants comme les notions d'identité et de relations à travers les dimensions spatiales et temporelles. Pour un temps donné, nous catégorisons les relations spatiales et de filiation et, à travers le temps, les filiations temporelles et les relations spatio-temporelles. Les structures de graphe émergentes permettent non seulement de caractériser l'évolution d'un ensemble d'entités spatiales, mais aussi de découvrir de nouvelles propriétés. Des fonctions de manipulation de graphe sont développées et appliquées au modèle de graphe spatio-temporel. Ces fonctions identifient des processus génériques (e.g. vie et mort d'une entité) ou liés à une application spécifique et à sa sémantique. Afin de combiner plusieurs sources d'informations au sein d'un même graphe, des fonctions de jointure permettent l'intégration de plusieurs graphes au sein d'une représentation unifiée. Les propriétés des graphes ainsi constitués et le rôle des différentes entités présentes au sein de ces graphes sont étudiés par une qualification des différents types de routes les reliant. La consistance du modèle de graphe spatio-temporel est abordée à partir d'un algorithme de vérification de contraintes. Nous faisons la différence entre les contraintes de domaine intrinsèques au modèle de graphe, et les contraintes sémantiques dépendantes d'une application en particulier. Une extension de la démarche de modélisation est réalisée à partir d'une structure basée sur les bigraphes qui permet de représenter explicitement un phénomène spatio-temporel selon plusieurs niveaux de granularité spatiale.

Mots-clés : modélisation spatio-temporelle, théorie des graphes, spécification de contraintes

A SPATIO-TEMPORAL GRAPH-BASED MODEL FOR THE EVOLUTION OF GEOGRAPHICAL ENTITIES

Abstract

Space and time are fundamental concepts needed for the analysis of the evolution of spatio-temporal entities. In order to model these evolutions, the changes and relations that modified and characterise entities must be identified. These changes are especially qualified by modifications of their position, spatial footprint, or their identity.

Spatio-temporal modelling requires the development of solutions of representation which allow for the identification and study of processes involved, of their causes and their consequences. The fact that a formal representation and computing of a phenomenon leads to a simplified representation of the reality, implies that this modelling should keep the level of details available and/or aimed at for this description. Ideally, it should be possible to study a phenomenon at several levels of granularity in space and time. If numerous approaches have been proposed for the modelling of the evolution of entities in spatio-temporal phenomena, none has a structure which fully takes into account the semantics of these applications.

This research proposes a spatio-temporal graph model which allows for the characterization of main properties of spatio-temporal evolution. Several concepts are distinguished, such as the notions of identity and relations, through the spatial and time dimensions. For a given time, we categorize spatial and filiation relations, and through time, temporal filiation and spatio-temporal relations. The emerging graph structures characterize the evolution of a set of spatial entities and permit to discover new properties. Graph manipulation functions are developed and applied to the spatio-temporal graph model. These functions identify generic processes (e.g. life and death of an entity) or processes linked to a specific application and its semantics. In order to combine several sources of information in a single graph, graph join functions permit the integration of several graphs in a unified representation. The properties of these graphs and the role of their entities are studied by a qualification of the different types of routes that link these entities. The consistency of the spatio-temporal graph model is verified by a constraints-checking algorithm. We make a difference between domain constraints belonging to the graph model, and semantic constraints which depend on a specific application. An extension of the modelling approach is achieved from a structure based on bigraphs, that facilitates the explicit representation of a phenomenon according to different levels of spatial granularity.

Keywords: spatio-temporal modelling, graph theory, constraints specification

Discipline : Informatique/Computer Science