

Lectures

6 janvier 2021

1 Notes de lectures

Papier	Catégorie	Notes
ON THE LOGIC OF THEORY CHANGE : PARTIAL MEET CONTRACTION AND REVISION FUNCTIONS [1]		Contraction : On rejette une proposition x qui était auparavant dans une théorie A . La difficulté est souvent de savoir quelles propositions doivent être rejetées en même temps que x pour assurer la théorie soit fermée sous conséquence logique. Révision : On ajoute à la théorie A une proposition x , inconsistente à la théorie, et on doit alors réviser la théorie pour retrouver sa consistance au vu de ce nouveau fait. Levi identity : if $A - x$ denotes the contraction of A by x , then the revision of A by x , denoted $A + x$, can be defined as $Cn((A - x) \cup x)$, where Cn is a given consequence operation. Définit des critères que doit valider le résultat d'un opérateur de révision sur une knowledge base.
An architecture of selective forgetting [9]		S'intéresse aux cas où les données sont liées entre elles. On en oublie une, que pasa pour celles liées ? - Pour les backwards références, c'est compliqué. Si on oublie un fait qui a été inféré d'autre chose, on risque juste de le ré-inférer à l'infini puisque les causes sont encore présentes. On part donc du principe qu'on va uniquement oublier des choses de la base de faits originale avant inférence (initial knowledge). - Pour les forward références, deux cas : consolidation (conséquence devient axiome), ou abstraction, qui revient à mettre la donnée oubliée dans un état particulier (true ou false).
On the difference between updating a knowledge base and revising it [11]		Update : le monde change, on change donc la base de connaissance. Ex : « le salaire de Joe augmente de 5Re- vision : On change qqc dans un monde qu'on pense statique. Exemple : un test qu'on pensait vrai s'avère faux. Une des grosses différences est qu'on considère pour la révision que le fait initial était une erreur à la base et n'aurait jamais dû être inclus dans la base de fait, alors que pour l'update c'est juste une évolution. Propose des postulats pour le résultat d'un update, différents (mais proches) de ceux de la révision.
Intentional Forgetting in Artificial Intelligence Systems : Perspectives and Challenges [24]		Programme « Intentional Forgetting in organizations » : paradigme interdisciplinaire entre informatique et psychologie. 8 projets, dont 5 orientés IA. Site du programme : http://www.spp1921.de/index.html.en
Psychological perspective on intentional forgetting [7]		Présente différentes théories de l'oubli, d'un point de vue psycho. Fait la distinction entre oubli intentionnel, et les autres formes d'oubli. Présente une vision à l'échelle de l'individu, ou des groupes, et fait un lien avec l'IA.
Experiments in cultural language evolution - Intro [21]		Donnes quelques propriétés du langage : expressive adequacy (être assez précis sur ce qu'on veut dire), cognitive effort (qu'il soit facile à dire, et à comprendre), learnability (possibilité de l'étendre à de nouveaux phénomènes e.g), social conformity (une certaine conformité pour permettre de se comprendre). Ces propriétés vont agir comme pression sélective pour l'émergence d'un langage. Présente le principe des "language games". Lire les autres chapitres pour avoir une meilleure idée des expériences.

Classification des mécanismes organisationnels dans les réseaux d'agents [15]		<p>Propose une classification des mécanismes organisationnels pour les SMA, et donne pour chacun plusieurs exemples. Classification proposée :</p> <pre> graph TD A[Mécanismes organisationnels] --> B[structuraux] A --> C[fonctionnels] B --> D[activation] B --> E[découverte] B --> F[situations] C --> G[délégation] C --> H[hiérarchiques] C --> I[répartition] </pre> <p>Peut être utile au moment de créer les jeux expérimentaux.</p>
Une architecture d'agent BDI basée sur la théorie des fonctions de croyance : application à la simulation du comportement des agriculteurs [22]		Bon rappel avec exemple sur le fonctionnement des agents BDI, avec un exemple concret.
Intentional forgetting in distributed AI [20]		<p>Dans les cas où il faut bcp (trop) d'infos pour traiter un pb, les SMA vont souvent distribuer le problème afin de réduire sa complexité. Deux cas d'information overload (IO) : dimension qualitative, trop d'info car environnement trop riche -> agent spécialisés. Dimension quantitative, problème de la correspondance des informations, les agents peuvent s'échanger des infos fausses ou obsolètes -> bonne coordination. Compromis à trouver pour savoir commun : en avoir permet une meilleure coordo et + de redondance, moins de risque de perte de savoir utile, mais trop entraîne des IO. Principe de "team cognition", quand le savoir est partagé entre agents spécialisés. Team coalition : agents avec des capacités hétérogènes. Team formation : capacités homogènes. Du coup leur Intentional Forgetting sert + à une réorganisation du savoir. Les agents utilisés sont des "discourse agents", une variante de BDI. Pour définir les agents, on a l'ensemble des actions, et pour chaque agent un sous ensemble représentant les actions qu'il est autorisé à utiliser en fonction de son rôle dans l'équipe. En plus de BDI, les agents ont aussi un ensemble de plans "capability", représentant l'ensemble des plans qu'ils sont capables de réaliser. Une "transactive memory" permet de savoir qui est capable de quoi. Elle est individuelle, mais dans un monde idéal, tout le monde a une vision exacte des capacités de chacun, c'est donc un savoir partagé. Un modèle mental individuel également permet de savoir qui a quelles intentions. De la même manière, dans un monde parfait, il y a consensus là-dessus, et donc ce savoir individuel devient un savoir partagé. Quand un agent apprend ou oublie qqc, il suffit de le faire savoir à tlm pour qu'ils mettent leur TM à jour.</p>
Description logics[2]		<p>Analogie (personnelle, pas dans le papier), Tbox = classe, Abox = objet, instance. Dans le papier, Tbox = schema, Abox = data. Tbox est appelée "definitorial" ssi elle ne contient que des définitions, et qu'il n'y a pas de cycle. ALC une DL très utilisée, qui donne les bases, et a connu des extensions (par exemple "number restriction" pour ajouter des nombres à des définitions). Fait une comparaison entre DL, logique de 1er ordre et logique modale. A continuer. A noter (pas dans le papier) : le concept de "nominals" sort un peu du cadre Abox/Tbox, et peut permettre de faire référence aux datas au sein de la Tbox (par exemple qqn de telle classe doit forcément appartenir à telle asso, telle asso ou telle asso, sachant que les assos en question font partie de la Abox).</p>
Towards Simulation-Based Role Optimization in Organizations [19]		Tente de résoudre le problème Job-Shop-Scheduling avec une approche multi-agent. Cherche avant tout à maximiser la répartition initiale des rôles, donc pas forcément utile pr moi, mais donne une bonne idée des objectifs de AdaptPro avec leur oubli intentionnel.

PR-OWL : A Bayesian Ontology Language for the Semantic Web [5]		Présente différentes alternatives bayésiennes. Les réseaux bayésiens classiques ne sont pas assez expressifs. L'approche proposée se base sur la logique MEBN (Multi Entity Bayesian Network), qui combine probas et logique de premier ordre. Certaines requêtes peuvent toutefois s'avérer indécidables. A continuer après + de recherches sur réseaux bayésiens
Bayesian Network [18]		Apprendre dans un tel réseau est un peu comme entrainer un NN : c'est faire varier les probas entre les entités (les poids des liens entre les noeuds). Un réseau causal est un réseau où les parents d'un noeud sont sa cause directe. Si on fixe la valeur d'un noeud, on retire alors le lien avec ses parents. Un peu vieux, donc probablement plus à jour, mais explique bien la base des réseaux bayésiens, et les idées derrière.
Advances in Bayesian network modelling : Integration of modelling technologies [16]		Présente de nombreux champs d'applications où des BN ont été utilisés. Peut être intéressant : agent based modeling. IBN, des BN intégrés à d'autres modèles. Ajd, ils sont surtout statiques (réalisés par des experts).
Bayesian or biased ? Analytic thinking and political belief updating [23]		Etude de psycho cognitive. On pose des questions orientées politiquement à des personnes. Après chaque réponse, un signal indique vrai/faux. Ils ont été informés que ce signal donnait la bonne réponse 2 fois sur 3. On s'intéresse alors à comment les gens ont update leurs croyances, en leur posant les mêmes questions une seconde fois. Intéressant, mais se focalise surtout sur des idées polémiques, et donc des raisonnements motivés, pas des cas plus basiques. Etablit une corrélation (faible, mais significative statistiquement) entre score au RCT et capacité à update ses croyances en restant proche d'un modèle bayésien. Le test RCT peut en revanche être à creuser, illustre la capacité des individus à mobiliser + ou - de connaissances en fonction du contexte, rappelle le modèle "système 1 système 2" de Kahnemann, p-e un lien à faire avec oubli temporaire ?
Bayesian models of cognition [14]		Cite au début bcp de modèles bayésiens utilisés en science cognitive, peut être utile. Un peu technique, mais présente différentes approches bayésiennes, dont les BNs, et leurs fondements mathématiques (surtout utile sur ce point).
Bayesian theories of conditioning in a changing world [4]		Papier qui cite des études montrant que chez les animaux, la surprise (par exemple, conditionner l'animal à lui donner de la bouffe après une sonnerie, mais au bout d'un moment lui envoyer un choc électrique), et donne une interprétation bayésienne à ce fait (en gros, surprise -> grosse incertitude sur leurs croyances actuelles -> plus gros updates dans les probas de leurs croyances). A creuser, voir si ce mécanisme existe aussi pr l'oubli ? (si on oublie bcp de chose, ça peut vouloir dire que l'envt a bcp changé, donc entrainer moins de confiance dans nos connaissances actuelles et rendre l'oubli + probable ?). Direction à creuser.
The case for Motivated Reasoning [?]		Présente le phénomène de "raisonnement motivé". Historiquement, une thèse concurrente était de dire que simplement, pour certains individus, certaines croyances étaient + plus plausibles au vu de leurs croyances initiales (prior believes) Peut être intéressant pour les oublis temporaires. (à continuer)
Revision in network of ontologies [8]		Commence par définir les propriétés d'un réseau d'ontologie. "Revision defines operators for modifying a theory (K) in the logic when some action occurs, like the knowledge that an assertion must be believed (revision), must not be believed (retraction), or does hold after a change (update). These operators are constrained by several postulates". Propose ensuite des opérateurs (dérivés de partial meet revision) pour de la révision dans un réseau d'ontologies.

The Role of Forgetting in Learning [16]		Dans le cadre de résolution de problèmes, on peut quantifier l'utilité d'un savoir : on résout le problème avec ou sans ce savoir, et la différence donne une idée sur la pertinence de ce savoir. Un savoir avec une valeur négative peut être oublié. Toutefois, un savoir faux peut s'avérer mieux que pas de savoir du tout : penser que tous les grizzlis sont des mangeurs d'humain (savoir trop généralisé) ; n'avoir aucune info sur leur dangerosité. Au contraire, des savoirs corrects peuvent entraîner à des erreurs (cite un papier). Mène ensuite des expériences sur de la recherche dans des graphes. Pas forcément hyper proche de ce que je fais, mais à garder en tête, peut-être creuser les aspects où au-delà d'être inutile, le savoir devient contre-productif, même si vrai.
Intentional Forgetting : An Emerging Field in AI and Beyond [3]		Fait un point sur l'oubli intentionnel dans l'IA en présentant différents papiers/projets (surtout en lien avec le programme de recherche allemand, mais pas que, biblio à creuser).
A Description Logic Primer [12]		Très bonne intro aux DL, clair. Ne pas hésiter à revenir dessus si besoin.
Processes of Successful Intentional Forgetting [10]		Papier de psycho cognitive, qui propose un framework pour l'intentional forgetting. Intentional forgetting is defined as a motivated attempt to limit the future expression of aspecific memory content. This definition differs from that of spontaneous forgetting, which occurs without motivation and regardless of the information's validity or relevance. Fait une distinction entre les études où les sujets sont prévenus ou non sur le fait qu'il faudra oublier : cette info changera la manière dont ils vont encoder l'info. Dans le cas où ils le savent, on ne considère pas ça comme intentional forgetting. Fait une différence entre deux approches de traitement de l'IF : les processus "memory-based", par exemple l'inhibition d'une donnée au sein de la mémoire, ou la ségrégation, le fait d'avoir une donnée en mémoire, mais avec un "flag" indiquant qu'elle ne doit pas être prise en compte. De l'autre côté, les processus decision-based, qui eux vont agir sur le R-cued, le signal de rappel d'une donnée, et vont faire que ce signal va moins probablement renvoyer la donnée oubliée. Présente un framework tentant de regrouper toutes ces approches (papier ancien, peut-être plus à jour, voir les recherches + récentes).
A Brief Survey on Forgetting from a Knowledge Representation and Reasoning Perspective [24]		Utilise une déf un peu différente de processus d'oubli intentionnel, par exemple ne pas prendre en compte toutes nos connaissances pour résoudre un problème particulier est considéré comme oubli intentionnel. Pour eux, abstraction = red whine and white whine merged to whine. Logique propositionnelle, oubli de variables : semantic changes should affect only the forgotten variable p, càd que le résultat d'une interprétation avant/après oubli doit tjrs être le même pour au moins une interprétation de la valeur oubliée. Définit logiquement un tel oubli. Variable forgetting aims at eliminating any occurrence of a variable. A more fine-grained extension considered by Lang et al. [60] is to forget literals, where an atom p and its negation not(p) are distinguished. Définit logiquement l'oubli de literals (en gros, oublier qu'une telle variable peut être vraie, ou fausse, et ne garder que les cas où elle est fausse, ou vraie, exemple 9 illustre bien). Logique de premier ordre, oubli de faits, bien illustré exemple 12. Oubli de relation : par exemple si on veut oublier "student", et qu'on a "student(Jon)", on va remplacer par "Il existe R tq R(John)", donnant une formule du second ordre (exemple 13). Enfin, oubli d'individu, via l'anonymisation : student(John) devient, après l'oubli de John, Il existe x tq student(x). Pour les DL, parle de la difficulté à oublier tout en conservant la similarité des modèles : c'est parfois impossible/indécidable. Cite de nombreuses refs de travaux en lien avec l'oubli en DL. Parle ensuite d'oubli pour ASP et logiques modales. Evoque certaines raisons de l'oubli, par exemple variables indépendantes, qu'on peut supprimer sans que cela n'ait d'impact sur la KB.

2 Champs libre, notes en vrac

Regarder les recherches du laboratoire de psychologie et neurocognition

Idées en vrac : Raisonnement motivé - cherry picking, variables indépendantes, système de Kahnemann...

Références

- [1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the Logic of Theory Change : Partial Meet Contraction and Revision Functions. *The Journal of Symbolic Logic*, 50(2) :510–530, 1985. Publisher : Association for Symbolic Logic.
- [2] Franz Baader, Ian Horrocks, and Ulrike Sattler. Chapter 3 Description Logics. In Frank van Harmelen, Vladimir Lifschitz, and Bruce Porter, editors, *Foundations of Artificial Intelligence*, volume 3 of *Handbook of Knowledge Representation*, pages 135–179. Elsevier, January 2008.
- [3] Christoph Beierle and Ingo J. Timm. Intentional Forgetting : An Emerging Field in AI and Beyond. *KI - Künstliche Intelligenz*, 33(1) :5–8, March 2019.
- [4] Aaron C. Courville, Nathaniel D. Daw, and David S. Touretzky. Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, 10(7) :294–300, July 2006.
- [5] Paulo Cesar G. da Costa, Kathryn B. Laskey, and Kenneth J. Laskey. PR-OWL : A Bayesian Ontology Language for the Semantic Web. In Paulo Cesar G. da Costa, Claudia d’Amato, Nicola Fanizzi, Kathryn B. Laskey, Kenneth J. Laskey, Thomas Lukasiewicz, Matthias Nickles, and Michael Pool, editors, *Uncertainty Reasoning for the Semantic Web I*, Lecture Notes in Computer Science, pages 88–107, Berlin, Heidelberg, 2008. Springer.
- [6] Thomas Eiter and Gabriele Kern-Isberner. A Brief Survey on Forgetting from a Knowledge Representation and Reasoning Perspective. *KI - Künstliche Intelligenz*, 33(1) :9–33, March 2019.
- [7] Thomas Ellwart and Annette Kluge. Psychological Perspectives on Intentional Forgetting : An Overview of Concepts and Literature. *KI - Künstliche Intelligenz*, 33(1) :79–84, March 2019.
- [8] Jérôme Euzenat. Revision in networks of ontologies. *Artificial Intelligence*, 228 :195–216, November 2015.
- [9] Jérôme Euzenat and Libero Maesano. An architecture for selective forgetting. In Luc Steels and Barbara Smith, editors, *AISB91*, pages 117–128, London, 1991. Springer.
- [10] Hollyn M. Johnson. Processes of successful intentional forgetting. *Psychological Bulletin*, 116(2) :274–292, 1994. Place : US Publisher : American Psychological Association.
- [11] Hirofumi Katsuno and Alberto Mendelzon. On the difference between updating a knowledge base and revising it. 1991.
- [12] Markus Krötzsch, Frantisek Simancik, and Ian Horrocks. A Description Logic Primer. *arXiv :1201.4089 [cs]*, June 2013. arXiv : 1201.4089.
- [13] Ziva Kunda. The case for motivated reasoning. *Psychological bulletin*, 108(3) :480, 1990. Publisher : American Psychological Association.
- [14] Thomas L. Griffiths, Charles Kemp, and Joshua B. Tenenbaum. Bayesian models of cognition. April 2008. Publisher : Carnegie Mellon University.
- [15] Laurent Lacomme, Yves Demazeau, and Valérie Camps. Classification des mécanismes organisationnels dans les réseaux d’agents. In *JFSMA*, pages 79–88, 2009.
- [16] Bruce G. Marcot and Trent D. Penman. Advances in Bayesian network modelling : Integration of modelling technologies. *Environmental Modelling & Software*, 111 :386–393, January 2019.
- [17] SHAUL Markovitch and PAUL D. Scott. The Role of Forgetting in Learning. In John Laird, editor, *Machine Learning Proceedings 1988*, pages 459–465. Morgan Kaufmann, San Francisco (CA), January 1988.
- [18] Judea Pearl. Bayesian networks. 2011.
- [19] Lukas Reuter, Jan Ole Berndt, and Ingo J. Timm. Towards Simulation-Based Role Optimization in Organizations. In Gabriele Kern-Isberner, Johannes Fürnkranz, and Matthias Thimm, editors, *KI 2017 : Advances in Artificial Intelligence*, Lecture Notes in Computer Science, pages 359–365, Cham, 2017. Springer International Publishing.
- [20] Lukas Reuter, Jan Ole Berndt, Anna-Sophie Ulfert, Conny H. Antoni, Thomas Ellwart, and Ingo J. Timm. Intentional Forgetting in Distributed Artificial Intelligence. *KI - Künstliche Intelligenz*, 33(1) :69–77, March 2019.
- [21] Luc Steels. Experiments in Cultural Language Evolution, 2012.
- [22] Patrick Taillandier, Olivier Therond, and Benoit Gaudou. Une architecture d’agent BDI basée sur la théorie des fonctions de croyance : application à la simulation du comportement des agriculteurs. In *Journées Francophones sur les Systèmes Multi-Agents*, Journées Francophones sur les Systèmes Multi-Agents, pages 107–116, Honfleur, France, 2012. Cepadué.
- [23] Ben M. Tappin, Gordon Pennycook, and David G. Rand. Bayesian or biased? Analytic thinking and political belief updating. *Cognition*, 204 :104375, November 2020.
- [24] Ingo J. Timm, Steffen Staab, Michael Siebers, Claudia Schon, Ute Schmid, Kai Sauerwald, Lukas Reuter, Marco Ragni, Claudia Niederée, Heiko Maus, Gabriele Kern-Isberner, Christian Jilek, Paulina Friemann, Thomas Eiter, Andreas Dengel, Hannah Dames, Tanja Bock, Jan Ole Berndt, and Christoph Beierle. Intentional Forgetting in Artificial Intelligence Systems : Perspectives and Challenges. In Frank Trollmann and Anni-Yasmin Turhan, editors, *KI 2018 : Advances in Artificial Intelligence*, Lecture Notes in Computer Science, pages 357–365, Cham, 2018. Springer International Publishing.