Properties of Module Notions and Atomic Decomposition (Extended Abstract)[*]

Robin Nolte and Thomas Schneider

Department of Mathematics and Computer Science, University of Bremen, Germany

In ontology development, modularity has received great attention in the past years; see, e.g., the LNCS monograph [27]. The non-standard reasoning tasks of extracting modules and of decomposing an ontology into modules have manifold applications in ontology reuse, versioning, debugging, and comprehension, as well as collaborative ontology development and automated reasoning optimization.

When extracting a single module from a TBox $\mathcal{T}$, that is, a subset $\mathcal{M}$ that can be used as a proxy for $\mathcal{T}$, it is crucial for all these scenarios that $\mathcal{M}$ encapsulates the knowledge from $\mathcal{T}$ about a certain topic, which is usually taken to be a set of terms, the *seed signature* Σ. This encapsulation is typically captured via the notion of *Σ-inseparability* [17,2,3,15], which generalizes that of a conservative extension [4,20,13]. However, depending on the application, the widely adopted requirement that $\mathcal{M}$ be Σ-inseparable from $\mathcal{T}$ is not always sufficient. For example, when importing $\mathcal{M}$ in place of $\mathcal{T}$ into an external TBox, $\mathcal{M}$ should even be Σ'-inseparable from $\mathcal{T}$, where Σ' is the union of Σ and the signature of $\mathcal{M}$—this property, called *self-containment* [18], ensures that $\mathcal{M}$ encapsulates the knowledge about *all of its own terms*, making $\mathcal{M}$ a suitable proxy for $\mathcal{T}$ w.r.t. those terms rather than just Σ. On the other hand, in scenarios such as optimization of debugging and explanation [26,14] or of reasoning [7,30], $\mathcal{M}$ should even preserve all ways to derive the knowledge about Σ (or Σ'), captured by the notions of $\mathcal{M}$ being weakly (strongly) depleting or justification-preserving [18,1,24].

Decomposition aims at computing the modular structure of a TBox—a representative subset of all modules together with their logical interactions. This structure can be used to better understand the TBox, aid its collaborative design, and optimize tool support [9,10]. Among the available techniques, *atomic decomposition (AD)* [11] stands out by its efficiency and genericness: the underlying algorithm is based on a linear number of module extractions, for a suitable module notion. Originally based on locality-based modules (LBMs) [8], the AD framework was recently shown to work with any module extraction function m that yields uniquely determined Σ-inseparable subsets of the input TBox which satisfy certain module properties, among them self-containment [10]. In recent years, AD has received increased attention [29,21,19]

A wide range of module extraction functions and module properties are known [8,16,28,23,12,1,24,6,5,19]. The functions differ in the properties they ensure and, for a given module extraction function m and property P, it is not always obvious whether m satisfies P. This is particularly so when m is based on some normal form: e.g., there may be various ways to recover a module of an arbitrary TBox $\mathcal{T}$ from a module of its normalization, thus violating uniqueness and suitability for AD. Sometimes only weak properties are known, and their strong counterparts have to be ensured via iteration [1].

The aim of this paper, which is an extended abstract of our KR 2021 paper [22], is to provide an axiomatic approach to systematize the wealth of existing module properties for module extraction functions guaranteeing Σ-inseparability. This knowledge enables us to examine whether module notions besides LBMs can be used safely with AD.

We briefly report on our main results. For this, let a *module extraction function (MEF)* be a (partial) function $m(\cdot,\cdot)$ that maps a signature Σ (a set of concept and role names) and a TBox $\mathcal{T}$ to a subset $\mathcal{M}$ of $\mathcal{T}$ that is Σ-inseparable from $\mathcal{T}$. For a property P of a single module, we say that an MEF *m satisfies P* if $m(\Sigma,\mathcal{T})$ satisfies P for all Σ and $\mathcal{T}$, e.g., an MEF is depleting if it only yields depleting modules.

First, we conduct a systematic study of the relationships between module properties and we find, amongst others, the following implications.

Theorem 1. *Let m be an MEF. The following implications hold.*
1. *If m is monotonic in $\mathcal{T}$, m is justification-preserving;*
2. *If m is justification-preserving, m is depleting;*
3. *If m is strongly justification-preserving, m is strongly depleting and self-contained.*

These interrelations yield rigorous, short proofs of properties satisfied by various module notions. For example, LBMs are obviously monotonic in $\mathcal{T}$, substantiating the 'folklore' assumption of them being justification-preserving [25,1].

Second, we generalize the iteration process underlying various MEFs, which ensures self-containment—either inherent to the module extraction process itself [8,16,12] or as an extension of a non-self-contained MEF [1,24].

Theorem 2. *For any MEF m, there is an iterative algorithm that computes a self-contained MEF m^+ in linear time with access to an oracle of m. If m is justification-preserving (depleting), m^+ is strongly justification-preserving (strongly depleting).*

Third, various MEFs depend on the input TBox being normalized [1,28,23]. For AD, however, we need to compute self-contained modules of the original TBox. Unfortunately, denormalizing modules that satisfy one of the 'strong' properties turns out to be hard:

Theorem 3. *Under mild assumptions, extracting non-trivial modules from consistent TBoxes is polynomial-time Turing-reducible to the following problem: Given a P-module $\mathcal{M}'$ of a normalized TBox $\mathcal{T}'$, compute a P-module $\mathcal{M}$ of the original, non-normalized TBox $\mathcal{T}$ with $P \in \{$self-contained, strongly justification-preserving, strongly depleting$\}$.*

Fortunately, as it is known that denormalization of justification-preserving and depleting modules is possible [1] (as a side result, we also report on an improvement on how to do this more easily), the strong variants can still be ensured using Theorem 2.

Fourth, given a MEF that is monotonic in Σ, the above results enable us to construct a 'repair' of it that is suitable for computing the AD of a general TBox:

Theorem 4. *If an MEF m is monotonic in Σ, we can extend m to an MEF m' that satisfies all properties required to compute the AD of any general TBox. In natural cases, this holds even if the domain of m is restricted to normalized TBoxes.*

Finally, we conduct a case study with an existing family of module notions based on normalization and Datalog reasoning [1], which suit a wide range of applications, and show both that they fit our framework and that they can be repaired for use in AD.

Future work should address several theoretical questions and compare AD based on repaired MEFs with AD based on LBMs [10]. As LBMs guarantee a very strict notion of inseparability, we hope to see improvements with, e.g., Datalog based modules [1], which in contrast can be tailored to specific use cases.

References

1. Armas Romero, A., Kaminski, M., Cuenca Grau, B., Horrocks, I.: Module extraction in expressive ontology languages via Datalog reasoning. J. Artif. Intell. Res. **55**(1), 499–564 (2016)
2. Botoeva, E., Konev, B., Lutz, C., Ryzhikov, V., Wolter, F., Zakharyaschev, M.: Inseparability and conservative extensions of description logic ontologies: A survey. In: Tutorial Lectures of RW'16. LNCS, vol. 9885, pp. 27–89. Springer-Verlag (2016)
3. Botoeva, E., Lutz, C., Ryzhikov, V., Wolter, F., Zakharyaschev, M.: Query inseparability for $\mathcal{ALC}$ ontologies. Artif. Intell. **272**, 1–51 (2019)
4. Byers, P., Pitt, D.: Conservative extensions: A cautionary note. EATCS-Bulletin **41** (1990)
5. Chen, J., Ludwig, M., Ma, Y., Walther, D.: Computing minimal projection modules for $\mathcal{ELH}^r$-terminologies. In: Proc. of JELIA'19. LNCS, vol. 11468, pp. 355–370. Springer-Verlag (2019)
6. Chen, J., Ludwig, M., Walther, D.: Computing minimal subsumption modules of ontologies. In: Proc. of GCAI'18. EPiC Series in Computing, vol. 55, pp. 41–53. EasyChair (2018)
7. Cuenca Grau, B., Halaschek-Wiener, C., Kazakov, Y., Suntisrivaraporn, B.: Incremental classification of description logics ontologies. J. Autom. Reason. **44**(4), 337–369 (2010)
8. Cuenca Grau, B., Horrocks, I., Kazakov, Y., Sattler, U.: Modular reuse of ontologies: Theory and practice. J. Artif. Intell. Res. **31**, 273–318 (2008)
9. Cuenca Grau, B., Parsia, B., Sirin, E., Kalyanpur, A.: Modularity and web ontologies. In: Proc. of KR'06. pp. 198–209. AAAI Press (2006)
10. Del Vescovo, C., Horridge, M., Parsia, B., Sattler, U., Schneider, T., Zhao, H.: Modular structures and atomic decomposition in ontologies. J. Artif. Intell. Res. **69**, 963–1021 (2020)

11. Del Vescovo, C., Parsia, B., Sattler, U., Schneider, T.: The modular structure of an ontology: Atomic decomposition. In: Proc. of IJCAI'11. pp. 2232–2237. IJCAI/AAAI (2011)
12. Gatens, W., Konev, B., Wolter, F.: Lower and upper approximations for depleting modules of description logic ontologies. In: Proc. of ECAI'14. Frontiers in Artificial Intelligence and Applications, vol. 263, pp. 345–350. IOS Press (2014)
13. Ghilardi, S., Lutz, C., Wolter, F.: Did I damage my ontology? A case for conservative extensions in description logics. In: Proc. of KR'06. pp. 187–197. AAAI Press (2006)
14. Horridge, M., Parsia, B., Sattler, U.: Laconic and precise justifications in OWL. In: Proc. of ISWC'08. LNCS, vol. 5318, pp. 323–338. Springer-Verlag (2008)
15. Jung, J.C., Lutz, C., Martel, M., Schneider, T.: Conservative extensions in horn description logics with inverse roles. J. Artif. Intell. Res. **68**, 365–411 (2020)
16. Konev, B., Lutz, C., Walther, D., Wolter, F.: Semantic modularity and module extraction in description logics. In: Proc. of ECAI'08. pp. 55–59 (2008)
17. Konev, B., Lutz, C., Walther, D., Wolter, F.: Formal properties of modularisation. In: Stuckenschmidt et al. [27], pp. 25–66
18. Kontchakov, R., Pulina, L., Sattler, U., Schneider, T., Selmer, P., Wolter, F., Zakharyaschev, M.: Minimal module extraction from DL-Lite ontologies using QBF solvers. In: Proc. IJCAI'09. pp. 836–841 (2009)
19. Koopmann, P., Chen, J.: Deductive module extraction for expressive description logics. In: Proc. IJCAI'20. pp. 1636–1643. ijcai.org (2020)
20. Maibaum, T.S.E.: Conservative extensions, interpretations between theories and all that! In: Proc. of CAAP/FASE'97. LNCS, vol. 1214, pp. 40–66. Springer-Verlag (1997)
21. Martín-Recuerda, F., Walther, D.: Fast modularisation and atomic decomposition of ontologies using axiom dependency hypergraphs. In: Proc. of ISWC'14, Part II. LNCS, vol. 8797, pp. 49–64. Springer (2014)
22. Nolte, R., Schneider, T.: Properties of module notions and atomic decomposition, accepted for KR (2021)
23. Nortjé, R., Britz, K., Meyer, T.: Reachability modules for the description logic $\mathcal{SRIQ}$. In: Proc. of LPAR'19. vol. 8312, pp. 636–652. Springer-Verlag (2013)
24. Peñaloza, R., Mencía, C., Ignatiev, A., Marques-Silva, J.: Lean kernels in description logics. In: Proc. of ESWC'17, Part I. LNCS, vol. 10249, pp. 518–533 (2017)
25. Sattler, U., Schneider, T., Zakharyaschev, M.: Which kind of module should I extract? In: Proc. of DL'09. CEUR Workshop Proceedings, vol. 477. CEUR-WS.org (2009)
26. Schlobach, S., Cornet, R.: Non-standard reasoning services for the debugging of description logic terminologies. In: Proc. of IJCAI'03. pp. 355–362. Morgan Kaufmann (2003)
27. Stuckenschmidt, H., Parent, C., Spaccapietra, S. (eds.): Modular Ontologies: Concepts, Theories and Techniques for Knowledge Modularization, LNCS, vol. 5445. Springer-Verlag (2009)
28. Suntisrivaraporn, B.: Module extraction and incremental classification: A pragmatic approach for $\mathcal{EL}^+$ ontologies. In: Proc. of ESWC'08. LNCS, vol. 5021, pp. 230–244. Springer-Verlag (2008)
29. Tsarkov, D.: Improved algorithms for module extraction and atomic decomposition. In: Proc. of DL'12. CEUR Workshop Proceedings, vol. 846. CEUR-WS.org (2012)
30. Zhao, H., Sattler, U., Parsia, B.: Avoiding subsumption tests during classification using the atomic decomposition. In: Simkus, M., Weddell, G.E. (eds.) Proc. of DL'19. CEUR Workshop Proceedings, vol. 2373. CEUR-WS.org (2019), `http://ceur-ws.org/Vol-2373/paper-54.pdf`