

オントロジー構築過程におけるロール概念の抽出と整理の方法*

3J-5

久保成毅

古崎晃司

來村徳信

溝口理一郎†

大阪大学産業科学研究所‡

1 はじめに

近年、オントロジー研究が盛んとなっているが、オントロジーの利用に関する研究に比べて、構築方法に関する研究は少ない。そこで我々は Activity-First Method (AFM) という構築方法論 [Mizoguchi95, Takaoka96, 久保 99] を提案し、AFM に基づいてオントロジー構築者 (オーサ) をガイド支援する計算機環境 [久保 98] の開発を進めている。本稿ではタスク・コンテキスト・ロールに関する構築方法について検討する。

2 ロール概念

2.1 タスク・ロール

オントロジーは問題解決の手順などに関するタスク知識を支えるタスク・オントロジーと問題解決の対象に関するドメイン知識を支えるドメイン・オントロジーとに分けて記述される [Mizoguchi95]。知識の共有・再利用の観点から、タスクとドメインは互いに依存性が低くなるよう定義する必要がある。

タスク・オントロジーを構成する主要な概念にタスク・アクティビティ (T-Act) とタスク・ロールがある。「プラントの復旧運転」(図 1) では「(復旧方策を) 決定する」、「実施する」といった概念が T-Act であり、問題解決の操作や処理を表す。T-Act はタスクの種類によらず共通に表れる概念が多く、またその他の概念は T-Act に伴って表れる。そこで AFM ではまずこの T-Act を抽出・整理し、その T-Act をもとに他の概念を抽出・整理する。T-Act は入力物や出力物といった対象を必ず伴う。この対象の実体はドメインのオブジェクトやその属性などであるが、問題解決の中で用いられる際には、問題解決過程での役割が与えられる。この役割を表す概念がタスク・ロールである。「プラント内温度を測定する」という処理では、「プラント内温度」はドメインの概念であるが、この処理中では「測定対象」という役割を担っている。タスク・ロールそのものはドメイン概念とは独立に、タスクの目的や構造によって定義される。タスク・ロールという概念によりタスクとドメインを結びつけることで、独立性が高く、かつ利用しやすい問題解決知識の定義を行うことができる。

タスク・ロールはタスクから見たロール概念であるが、一般にロール概念 [溝口 99] は対象への視点を反映する。そのため、同じ実体に対して視点の異なる複数のロールが与えられたり、問題解決中の処理の推移に伴いロールが変化したりする。本稿では T-Act の入出力物のタスク・ロールについて検討する。

2.2 タスク・コンテキスト・ロール

T-Act の入出力物のタスク・ロールにはタスク・コンテキスト・ロール (TC-Role) とアクティビティ・ロール (A-Role) がある。TC-Role と A-Role ではグレイン・サイズが異なる。TC-Role はそのタスク全体の中での役割であり、A-Role は実行されている T-Act から決まる役割である。復旧運転 (第 3

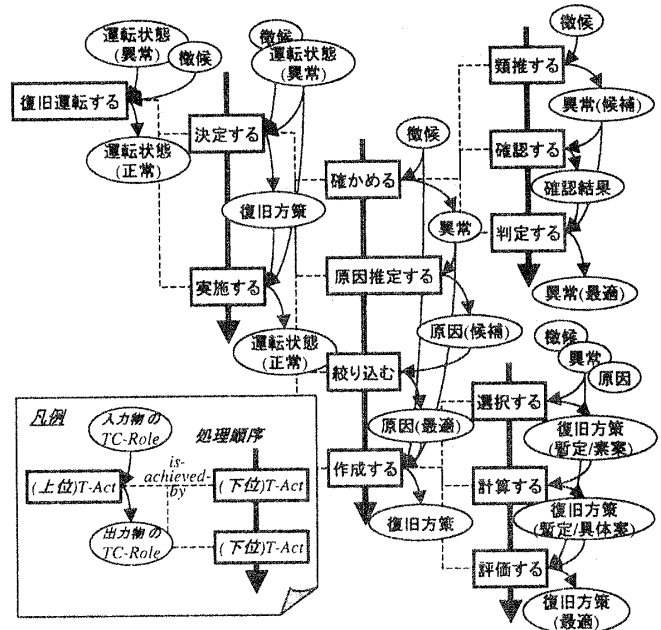


図 1: 復旧運転タスク is-achieved-by 階層と TC-Role

階層) では「(原因候補から原因を) 絞り込む」という処理がある。この入力物の A-Role は「絞り込む」のみから決まり単に「候補」と言われるが、復旧運転タスクの中では「異常状態を引き起こした原因の候補」という役割を持つので、TC-Role は「原因候補」となる。このように TC-Role はタスク・コンテキストと強く結合したロール概念であるのに対し、A-Role はタスク・コンテキストへの依存性が低く、タスク・コンテキストから独立した形で定義でき、「絞り込む」の入力物の A-Role は復旧運転以外のタスクで用いられても「候補」と呼ばれることが多い。

2.3 TC-Role の分類

T-Act はより小さなグレイン・サイズの T-Act の組み合わせで達成される。「絞り込む」は「(原因候補に伴って表れる異常の候補を) 予測し」、「(異常候補を) 確認し」、最後に「(原因候補が妥当かを) 判定する」という 3 つの T-Act から成る。このとき、「絞り込む」と 3 つの T-Act の間には「is-achieved-by」関係が成り立つと言う。このような視点から T-Act を整理したものが is-achieved-by 階層で、タスクの構造を明示化する。以下では is-achieved-by の上位、下位を単に上位、下位と呼ぶ。

TC-Role は上位-下位関係や順序関係などで決まる T-Act の目的 (タスク・コンテキスト) に依存して定義される。この T-Act と TC-Role の関連性から TC-Role を分類できる。

- 最上位 T-Act 入出力物の TC-Role: 最上位 T-Act の入出力物の TC-Role はその T-Act にしか依存しないため、ほぼその T-Act の A-Role と同じ定義となる。
- 上位入出力物と一致する物の TC-Role: 下位 T-Act は上位 T-Act の内部処理を明示化したものと捉えることができる。そのため目的は共有されており、上位 T-Act の入出力

*A method for articulation of role concepts in ontologies

†Shigeki Kubo, Kouji Kozaki, Yoshinobu Kitamura, Richiro Mizoguchi ‡I.S.I.R., Osaka Univ (Mizoguchi-lab) 〒567-0047 8-1, Mihogaoka, Ibaraki, Osaka Tel: 06-6879-8416 Fax: 06-6879-2123 e-mail: shigeki@ei.sanken.osaka-u.ac.jp

物と一致する物は、TC-Role も一致する。第 3 階層の「確かめる」、第 4 階層の「類推する」などの入力物は、第 2 階層の「決定する」の入力物と一致し、その TC-Role は「徴候」となる。

●内部処理生成物の TC-Role: 上位 T-Act の入出力物と一致しない物は、その T-Act の内部処理で必要とされる物である。これに与えられる TC-Role は次のような依存関係がある。

- 上位 T-Act 依存: 「実施する」の入力物「復旧方策」は上位 T-Act 「復旧運転する」に依存し、「復旧運転する」の内部処理に表れる TC-Role であることが記述される。
- 支配的 T-Act 依存: 「復旧方策」は、「決定する」の入力物でもあるが、この一連の処理では「実施する」が特に重要な T-Act であるため、この T-Act に依存した形で定義される。なお、このとき「実施する」は「決定する」に対して“支配的”である、と表現される。

実際には明確にどちらかに分類できないことが多く、「復旧(上位 T-Act「復旧運転する」に依存)方策(支配的 T-Act「実施する」に依存)」のようにその両方に依存している場合もある。多くの場合、上位 T-Act を考慮して、支配的な T-Act の入出力物の A-Role に意味制約が加えたものが TC-Role となる。

●状態 Role: 状態 Role は処理の進行に伴う状態の変化を表すロールである。復旧運転では「試作する」により「復旧方策暫定案」を作成し、「評価する」を経て「復旧方策最適案」となる。この「暫定案/最適案」というタスク・ロールは、処理の過程で同じアイデンティティを持つ「復旧方策」の状態遷移[林 98]を表している。「復旧方策」に伴う「暫定案/最適案」の様に、他のタスク・ロールを補助する役割であることが多い。このような補助的役割の持つロールを本稿では「復旧方策(暫定)」のように () を用いて表す。状態 Role はその状態の変化を引き起こす T-Act に依存し、状態遷移が表れる階層でのみ意識される。「復旧方策(最適)」の「最適」は「暫定」→「最適」という状態変化を引き起こした「評価する」に依存する。また、第 3 階層の「作成する」では単に「復旧方策」として意識される。状態 Role は支配的でない T-Act などの A-Role が上記の TC-Role の付加情報として表れたものであることが多い。

3 TC-Role 抽出・整理手順 (AFM)

AFM では次の過程で概念の抽出・整理を進める。現在 AFM では問題解決知識を持つ専門家が記述した、主にタスク・プロセスに関するドキュメントからオントロジーを構築することを想定して。なお、ステップ 1-3 は計算機による支援環境のプロトタイプが実装されている。

1. **T-Act 抽出:** ソースドキュメントから T-Act を抽出する。
2. **T-Act 整理:** T-Act を整理し、is-a 階層および is-achieved-by 階層を構築する。
3. **T-Act 定義/A-Role 抽出:** T-Act について、その入出力物を中心に定義する。同時に、入出力物の A-Role を定義する。
4. **オブジェクト・フロー構築:** タスク処理過程でのオブジェクトの流れを明示化したものをオブジェクト・フローと呼ぶ。各処理 (T-Act) の入出力物は前ステップで定義され、また処理の順序は T-Act の is-achieved-by 階層などにより明らかになっているので、それらをもとに、出力物と入力物のつながりを明示化していく。例えば (図 2), 前ステップまでの作業の結果、「類推する」の出力物 (A-Role:類推結果) と「判定する」の入力物 (A-Role:判定対象) とが分かっている。ここで、この二つが同一物であることを明示化し、タスク内でのオブジェクトの流れを明らかにする。
5. **TC-Role 定義:** is-achieved-by 最上位の T-Act の入出力物から始めて、順次下位の T-Act へ作業を進める。まず最初

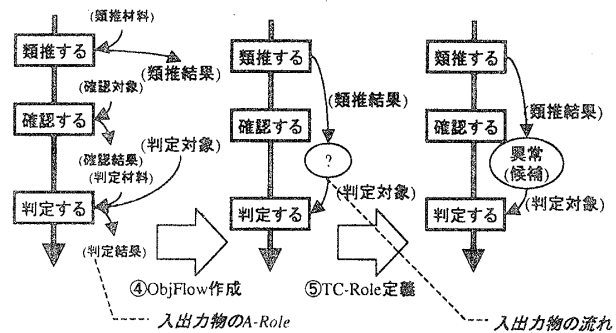


図 2: ObjFlow 構築と TC-Role 定義

に図 1 の第 1 階層の「徴候」、「運転状態」が定義される。以下第 3 階層を例に説明する。

- (i) 第 2 階層の「決定する」の入出力物と一致する「徴候」、「復旧方策」などが定まる。
- (ii) 「原因推定する」は「確かめる」、「絞り込む」に対して支配的である。「復旧運転する」を考慮して「原因推定する」の入出力物の TC-Role が「異常」、「原因」と定まる。
- (iii) 「絞り込む」は「原因推定する」を補助する T-Act であるので、その入出力物は状態変化のみとして捉え、「候補」、「最適」が定まる。

TC-Role を定義する際には、どの T-Act に依存しているのかを把握する必要がある。この点に関して、計算機による履歴管理などの支援が可能と考えている。

6. **ドメイン概念の抽出・整理・定義:** ドメイン概念は膨大であるが、問題解決で必要とされる概念は、その問題解決中の処理 (T-Act) の対象となるものである。そこで、既に抽出した TC-Role と結びついているドメイン概念を中心に抽出を行う。このステップでは、計算機によりソースドキュメントと T-Act の対応関係を管理することで、TC-Role と結びつくドメイン概念の抽出の労力を軽減する支援を行う予定である。

4 まとめ

タスクとドメインの独立性と利用容易性を両立させるためにはタスク・ロールの導入が必要である。本稿では問題解決の処理過程での入出力物のロールを表すタスク・コンテキスト・ロールがどのように抽出・定義されるかについて検討した。

今後、抽出・定義された TC-Role をもとにドメイン概念の抽出・整理を行う方法について検討を進める予定である。

参考文献

- [林 98] 林雄介, 他: 「概念間関係に関するオントロジー的考察～ is-a, part-of, identity ～」: 電子情報通信学会技術研究報告 Vol.98 No.436 pp1-8(1998)
- [久保 98] 久保成毅, 他: 「オントロジー構築ガイドシステム の概念設計」: 第 12 回人工知能学会全国大会論文集 pp36-39(1998)
- [久保 99] 久保成毅, 他: 「オントロジー構築方法 AFM(Activity-First Method) の詳細化の試み」: 第 13 回人工知能学会全国大会論文集 pp114-117(1999)
- [Mizoguchi95] Riichiro Mizoguchi, et al: "Ontology for Modeling the World from Problem Solving Perspectives": IJCAI Workshop on Basic Ontological Issues in Knowledge Sharing(1995)
- [溝口 99] 溝口理一郎, 他: 「オントロジー-工学基礎論-意味リンク, クラス, 関係, ロールのオントロジーの意味論」: 人工知能学会誌 Vol.14 No.6 掲載予定 (1999)
- [Takaoka96] Y. Takaoka, et al: Identification of Ontologies to Reuse Knowledge for Substation Fault Recovery Support System: Decision Support Systems 18, pp.3-21, Elsevier(1996)