

OSI 適合性試験項目の複雑度による評価

若杉忠男

若杉情報技術コンサルタントオフィス

〒251 藤沢市片瀬山 3-1-11、Tel 0466-23-4832

OSI 適合性試験の国際的な調和のために試験スイートの質の評価と比較は重要である。試験項目をその必要度と複雑度の2つの尺度で評価するという方法があるが、本論文ではそのうちの複雑度についての客観的な定義を試みる。まずISO 9646で定義された4つの抽象試験法と、実際の試験システムを参考にして、R法とD法の中間のR*法、D法とC法を兼ね備えた(C+D)法という試験法を定義する。またこれらの試験法によって実施できる試験項目のグループをそれぞれ、R*、(C+D)などと表現する。そうすると、 $R^* \subseteq (C+D) \subseteq L$ という関係がある。これらのグループの内、より大きなグループには含まれているが小さなグループには含まれていない試験項目は、複雑な試験項目であると言える。この定義を実際に使用されているトランスポートの試験スイートに適用し、どのような試験項目がどのグループに分類されるかを示して、この分類が複雑度の評価として適当であることを示す。

Estimation of quality of OSI conformance test cases by complexity

TADAO WAKASUGI

WAKASUGI Information Technology Consultant Office
3-11-1 Kataseyama Fujisawa-city, 251 JAPAN

In estimation of quality of OSI conformance test suites, there are two aspects :necessity and complexity. In order to estimate complexity of test cases, we use the abstract test methods written in ISO 9646, and define new abstract test methods, R*-method and (C+D)-method. Among these methods, there is a relation " $R^* \subseteq (C+D) \subseteq L$ ", where \subseteq means that the test cases which can be applied to the left side of \subseteq can be applied to the right side. By this relations, we can classify test cases, and define complexity of them.

0. まえがき

OSI 適合性試験の問題の一つに試験の質の評価がある。すなわち、現在実施されている適合性試験は適合性の評価として充分か、偏りや足りないところがないか、また他の国で行われている試験と比較して質が同じ程度か、というような問題である。また現実に適合性試験を実施している試験センタにとっては、質のよい試験項目を選択して試験の効率化をはかりたいという要望もある。

試験項目の評価については多くの方法が研究されて

いる。たとえば、故意に埋め込んだフォールトのうち何%を発見できるかによってフォールトカバレッジを定義するとか[1]、試験項目を基本的なボタンやグループに分け限定された範囲内でのカバレッジを評価する[2] どのような方法がある。また、IEEEの標準推進団体であるPOSIXでは試験要素を試験レベルとComplexityレベルという2次元マトリックスに分類して評価している[3]。この試験レベルとComplexityレベルいう二つの評価基準は、たとえば売値と原価のようなもので、試験スイートの評価尺度として妥当なもの

と思われる。こういう観点からの評価については、筆者も試験項目の必要度と複雑度（発表時には試験実施困難度という言葉を使用した）によって試験項目を表1に示すようなマトリックスに分類し、試験項目の実施優先度を定めることを提案した[4]。

表1 試験項目の必要度、複雑度と実施優先順位

	高 ← 必要度 → 低						
低 ↑ 複雑度 ↓ 高	<table border="1"> <tr> <td>必須な試験項目</td> <td></td> </tr> <tr> <td>推薦する試験項目</td> <td></td> </tr> <tr> <td>実施は任意試験項目</td> <td></td> </tr> </table>	必須な試験項目		推薦する試験項目		実施は任意試験項目	
必須な試験項目							
推薦する試験項目							
実施は任意試験項目							

本論文では複雑度について、ISO 9646の抽象試験法に基づいて試験項目を分類することにより、複雑度の客観的な評価尺度を設定することを試みる。

1. 試験項目の複雑度

試験項目の開発方針はISO 9646 [5]に記述されているが、それによれば、プロトコルに記述された要件をすべて試験をして確認しなければならない。しかし一方では、[適合性試験は相互接続性を保証するものではない]ということが明記されているので、完全を求めているとはいえない。また、9646に書かれている抽象試験法には、R法、D法、C法、L法の4つがあるが、特にR法は、制御観測点（PCO）が試験対象（IUT）の下部にしかなく試験判定能力は限られたもので[6][7]、これを試験手法の一つとして認めたということは、結果的に100%網羅する試験を要求していないことになる。

IEEEの適合性試験基準[3]では、試験項目にあたる概念をアサーションという言葉を使って次のように分類をしている。

- ・基本アサーション：試験しなければならないもの
- ・拡張アサーション：基本アサーションでないもの

ここで拡張アサーションの条件はいくつかあるが、たとえば次のようなものである。

(1)このアサーションに対してはリライアブルな試験方法がない。

- (2)このアサーションの試験にはユーザによる非現実的な量のセットアップ手順が必要である。
- (3)このアサーション試験の作成には、非現実的な量試験開発時間が必要である。

また拡張アサーションについては、次のような試験結果コードを使用できるとしている。

PASS/FAIL/UNTESTED/UNRESOLVED

ここにUNTESTEDというコードがあるということは、現時点で実施困難な試験項目は技術が進歩するまで当面は実施しなくても仕方がないと考えていると解釈してよいであろう。

試験項目の複雑度に関連する尺度を具体的に定義した例としては、トランジションの長さによるものがある。すなわち、試験項目をFSMのトランジションで表現し、トランジションの長いものほど作成や実施に手間がかかると見ると、同じフォールトを発見するのに試験項目のトランジションの数が少ないほどよい[8][9]。したがって、テストビリティという尺度は、与えられたフォールト領域内で全てのフォールトをカバーできる最も少ないトランジションの長さの逆数と定義される。

筆者もこの考えにならって試験目的のトランジションの数による複雑度の定義を検討した。[8]では試験実施シーケンス全体の長さを考えているが、筆者の場合には試験目的のテストボディのトランジションを数えた。しかしこれには次のような問題がある。

- (1)試験手法（たとえばR法かD法かなど）によってまた判定にどの程度の信頼性を要求するか（遷移先の状態が確かに意図したものを確認するかどうかなど）によってトランジションの数が変わる。
- (2)テストボディとはどこから始まりどこで終わりとるかによって、トランジションの数が変わる。

ここでは簡単に、試験目的を表現する文体により分類することを試みる。文体が複雑な場合には試験項目を構成するトランジションの数も多く、したがって試験項目は複雑であると見るわけである。ここで①②③という記号を定義する。

①：[IUTが……する]という単文で記述できるもので、1トランジションに相当する。

例：Verify that IUT can accept a CC TPDU requesting TPDU size 128.

②：[もし…すると、IUTが…を…する]
という条件文で、一般にIFやWHENなどを含
み、2トランジションに相当する。

例：Verify That the IUT selects TPDU size 128
if receiving a CR TPDU proposing TPDU size
128. (if receiving を on receipt に変えて
も②とする)

③：複数の①や②を組み合わせた複合文、または[
反復][繰り返し]という言葉を使った文。AND
やTHENがよく使われ、3トランジション以上
になると想定されるもの。ただし試験目的にコメント
として文をつけたようなものは除く。

例：IUT establishes a connection with credit
N. LT sends N DT TPDUs, check that the IUT
responds with AK TPDUs.

この方法にも問題点はいくつかある。たとえば、上
記3つの文体にあてはめにくい文体がある。しかしそ
のような試験目的は少数であり、そのような場合には
上記のどれかの文体で表現できるかを調べるか、別途
トランジションを数えることにする。

このような文体による分類は分かりやすかつ試験
法によって変わるというようなことを配慮する必要が
ないというメリットがある。

本論文では、試験目的をまず適用できる試験手法に
よって分類し、さらにその中の一部について文体によ
る細分するという方法をとる。

2. 試験法のバリエーション

ISO9646では、R法、D法、C法、L法とい
う4つのタイプの抽象的試験法を定義しているが、こ
の4つの抽象試験法は、主としてPCOの位置によ
って分類される。PCOの位置にはIUTのすぐ上の上
位PCOとIUTの下位のPCOとがあるが、二つの
PCOをつなぎ情報をやりとりする仕掛として、C法
ではTMP (Test Management Protocol) がある。

しかし、実際の試験システムには、オペレータの協
力によりD法に近い機能をはたすR法とか、D法の条
件を満たすと共にC法の条件も満たすというシステム
とかがある。また、同じ試験目的でも、試験判定の確
実さをどの程度まで要求するかによって実施する試験
手法が異なる。たとえば、試験の結果、下位PCOで
異常が見つからなければ合格とするというならばR法
でよいが、上位PCOの結果もチェックするというな

らばD法が必要であり、さらに上位と下位の観測結果
の相互関係も確認すべきであると考えればC法が
必要になる[10]。

ここで抽象的試験法を複雑度の尺度として採用する
ために、ISO9646で定義された4つの抽象的試験
法と“現実”の試験システムとの違いを次に考察す
る。

2.1 R法とR*法

D法は上下にPCOを持つ。一方、R法は下位のP
COだけで上位PCOのない試験法である。したが
ってその判定の確実さも低い。上位PCO機能に相当す
るものが全くないとすると限られた試験しかできない
しかし実際のR法、たとえば情報処理相互運用技術協
会のテストなどを見ると、IUT側でのオペレータの
介入などによってD法とほぼ同じ内容の試験を実施し
ている。たとえばトランスポートでは簡単な上位テ
スタを使用し、FTAMでは上位にあるアプリケーション
プログラムを使用して、次のようなことを行っている。

- ・トランスポートの場合は簡単な上位テスタを使って
コネクションの確立要求、送信、切断要求などを実
施し、FTAMの場合はアプリケーションを使って
ファイル転送システムの初期化、読みだし/書き込
み、キャンセルなどを起動することが出来る。
- ・IUTが受信した情報を受け取り、予め用意してい
たものと同じかどうかチェックすることができる。
- ・クライアント側オペレータが電話連絡により、下位
テスタ側と作業の同期をとることができる。

ただし試験判定の精度は高いとは言えない。たと
えば、タイマの試験ではタイマの動く時間を正確には測
定できず、十分大きな時間の後に、コネクションが切
断されたことを確認してタイマが正しく作動したと考
える。

このような試験方法を、ここではR法とD法の間
にあるものとして、R*法と記述する。このR*法で
実施可能な試験項目の件数は相当に多いので、それに
先に述べた文体によるサブグループ化を組み合わせる
すなわち、R*法のうち、試験目的が単文からなるも
の(R*①と記述する)、条件文も含むもの(R*②
、さらに条件文の他に複合文も含むもの(R*③)
の3種である。そうするとこれらの間に次の式が成り
立つ。

$$R^*① \subseteq R^*② \subseteq R^*③ \subseteq D \quad (1)$$

2. 2 C法と(C+D)法

D法は上下にPCOを持つ。一方C法は下位PCOとIUTとが、規格化されたTMP (Test Management Protocol) によって規定され制御されるもので、上位PCOについては持つことを義務づけられていない。

TMPについては[規格化されたTMPは、IUTの最上位の抽象的サービス境界のすぐ上に位置する。IUTはUTと情報をやりとりするだけで解釈しない。またASPを提示できるか否かという点には左右されない。これはPCOではないからである][5]とされている。トランスポートのTMPの仕様書[11]にも、[TMPは、IUTに期待される結果を記述して送信するが、サービスプリミティブの存在を義務づけるものではない]という表現をしている。しかし実際にISOで作成されたC法によるトランスポート試験スイート[12]の内容は上位PCOがあるかのように開発されている。したがって本論文では上位PCOをもつC法、すなわちC法とD法の機能をもつテストを想定し、(C+D)法と記述する。そうするとD法で使用される試験項目は(C+D)法でも使用できる。これは次のように表される。

$$D \subseteq (C+D) \quad (2)$$

2. 3 L法

L法では、上位テストとIUTとの情報のやり取りは、標準化されたハードウェア装置で行われる。上位と下位のテスト間の連絡は同じシステム内で実施されるので時間遅れの影響も少ない。したがって(C+D)法よりも実施できる試験項目は多い。

3. 試験項目の複雑度による分類

前章で述べた式から次のような関係が得られる。

$$R^*① \subseteq R^*② \subseteq R^*③ \subseteq (C+D) \subseteq L \quad (3)$$

(3)式の右側の試験スイートは左の試験試験スイートを含む。この式を複雑度のレベルを表す式とみなし、右のグループに含まれ左には含まれない試験項目は、左に含まれるものより複雑であると定義する。

本章ではこの定義に基づいて具体的にトランスポートの抽象試験項目の分類を試み、この複雑度の定義の

妥当性を考察する。対象としては、ISOで定義したトランスポート[12]のクラス2を使う。この試験スイートは、本論文の(C+D)法に相当する。

(1) R*①

ここに入る試験項目は、全体の75%前後になる。すなわち次のようなものである。

- ・ 確立、送信、受信、切断などの基本的な試験項目
- ・ 無効な試験項目とInopportuneな試験項目はほとんどここに入る。

(2) R*②

- ・ CRとCC、EDとEA、DRとDCなど、二つで一組になる試験項目
- ・ ウィンドウ制御試験の単純な場合

(3) R*③

- ・ TSDUの分割・組立、連結・分離など、内容によってトランジションの数がいくつになるか不明な試験。
- ・ ウィンドウ制御や多重化などの、複雑な条件を組み合わせた試験
- ・ コネクションを切断した再確立するか、何種類ものデータを組合せて送信するというような反復繰り返し試験。

(4) (C+D)

- ・ 上位層からの不正要求/指示に対するIUTの反応を見る試験
- ・ R*法では、上位層機能を使って不正なデータを作成して入力するという試験は困難である。
- ・ 上位下位テストの同期をとる必要がある試験
- ・ 細かい条件の組合せが必要で、ヒューマンオペレータなどでは手に余るものとか、判定の確実さが欲しいもの。タイマの試験やウィンドウ制御など。

(5) L

- ・ 外部要因、特に上位と下位のPCOの相互連絡の影響を受けると実施/判定が困難な事象の試験。すなわち、タイミング/タイマのテストなどで、(C+D)法よりも判定精度が必要な試験。ただし下位層であるサービスプロバイダによる時間遅れの影響は防げない。

(6) L法でも実施できない／しない試験項目

いままで述べた試験項目の他に、L法でも試験できないが実際にシステム開発をする場合には必要という試験項目がある。

実装の開発にたずさわっている技術者の評価によると、R*法の試験項目は接続試験としてはほぼ満足でも、システム開発試験項目に較べると30%程度しかカバーしていないということである。これをここではシステム試験項目と呼ぶことにする。たとえば、次のようなものである。

(1) 下位層の影響を受ける試験。

- ・ タイマの正確さの確認や下位プロバイダの非同期動作の影響をチェックする試験

(2) ブラックボックス試験では技術的にできないもの

- ・ 外部から制御できないIUTの機能や、判定には必要であるがPCOからでは観測できない内部状態。ウィンドウ制御の場合の内部状態など。

(3) ISO9646のルールで適合性試験の範囲外とされているもの

- ・ プロファイルでは許されず、基本標準では許されている機能を使用していない事を確認する試験
- ・ IUTが起動側の場合、プロファイルでノットアプリ

カブルと決まっても、UTでその使用を抑制できない機能を、IUTが使用していても不合格としない。

- ・ ISO9646では、頑強さ(Robustness)試験、性能、信頼性、リアルイフェクト(Real Effect)試験などはOSIの適合性試験の範囲外であるとしている。
- ・ 製品システム固有の試験項目や、開発中にバグを見つけるために必要な試験、あるいは適合性解決試験と呼ばれている規格化されていない試験項目など。

以上の分類により複雑度を定義し、それによる試験項目の例を添えて表2に示す。また、ISOのトランスポート試験スイートについての各グループ別の試験項目件数を、表3に示す。

結局のところ、ISOのトランスポート試験スイートの試験項目のうち、タイマの試験に関する2件以外はすべてR*法で実施できるという結果になった。しかし判定精度に対する要求をあげれば、R*法の件数は減少して(C+D)法の件数が増える。

表2 試験項目の複雑度による分類と試験項目の例

分類	条件		試験項目例
R*①	上位のPCOはないがそれをオペレータがカバーするもの	単文で表現できる	基本的な1トランジション試験、特に無効試験はほとんどこれである
R*②		条件文で表現できるものも含む	上記の試験、及びネゴシエーションの試験
R*③		複合文で表現できるものも含む	上記試験、及び反復動作、多重化、フロー制御、連結・分離、結合・分解など
C+D	上下位からの制御観測可能かつPCOの間にTMPがある		上記試験、及び上位層からの不正要求試験、同期を必要とするタイマ試験
L	上位テストと下位テストの間に時間遅れがほとんどない		上記試験項目を含み、ブラックボックス試験のほとんどが実施できる
システム試験	システム開発時にメーカーとして施す試験		上記+IUTの内部状態を制御する試験項目、適合性試験の範囲外の試験項目

表3 各試験手法の試験項目の件数

複雑度	件数
R*①	206*****
R*②	247*****
R*③	273*****
(C+D)	275*****

4. まとめと今後の課題

OSIの適合性試験の実施方針はISO9646に定められているが、そのとおりに実施するのは現状では技術的にも経済的にも無理がある。そのために試験スイートの質を評価し、試験の質を落とさずに試験項目の件数を減らすことが必要である。ここでは複雑度という視点から客観的に分類し整理することを試みた。

そのために、実際に使われている適合性試験手法のタイプとしてR*法、(C+D)法というタイプを考えた。これによる試験項目の分類を、トランスポートの試験スイートに適用して分析を試みたが、C法やD法により試験項目が範囲が拡大することはほとんどなく、これらの方法のメリットは試験判定精度の向上と試験の手間の削減効果にあると言える。

関連テーマである試験項目の必要度の評価については別途発表する予定である。必要度は試験スイートひとまとめにして評価するが、複雑度については個々の試験項目ごとに評価した。今後の課題は、複雑度の考えを他のプロトコルにも適用して評価の妥当性を調べ、また、試験項目の必要度と組み合わせて評価して行く予定である。

謝辞

本論文は、東大齊藤忠夫教授、浅野正一郎教授、東北大白鳥則郎教授、湘南工科大杉山宏教授などをはじめ多くの方々のご指導とはげまして、さらに情報処理相互運用技術協会の資料提供によってまとめることができました。ここに心より感謝いたします。

参考文献

[1] G. v. Bochmann, A. Das, R. Dssouli, M. Dubuc, A. Ghedamsi, G. Luo: "Fault Models in Testing", Proc. of IWPTS#4, pp. II 19-32, (Oct. 1991).
 [2] R. Castanet, P. Corvisier, R. Casadessus: "In-

dustrial Experience on Test Suite Coverage Measurement", Proc. of IWPTS#5, pp.149-154, (Sep. 1992).

[3] "IEEE Standard Information Technology Test Methods for Measuring Conformance to POSIX", pp9-11, IEEE 1003.3, (March 21 1991).
 [4] T. Wakasugi: "A Study of OSI Conformance Tests - In the Case of Transport Layer", Proc. of IWPTS#3, (Oct. 1990).
 [5] ISO/IEC JTC 1/SC 21: "Information Technology - Open System Interconnection - Conformance Testing Methodology and Framework - Part 1/2/4/5", ISO 9646-1/2/4/5, (1994-3-14).
 [6] R. Dssouli, G. v. Bochmann: "Conformance Testing with Multiple Observers", Proc. of #6 of IFIP Workshop on Protocol Specification, Testing & verification, pp. 6.35-6.56, (June 1986).
 [7] G. Luo, R. Dssouli, G. v. Bochmann: "Generating Synchronizable Test Sequences based on Finite State Machine with Distributed Ports", Proc. of IWPTS #6, pp.143-158, (Sep. 1993).
 [8] A. Petrenko, R. Dssouli, H. Koenig: "On Evaluation of Testability of Protocol Structures", Proc. of IWPTS#6, pp. 114-127, (Sep. 1993).
 [9] R. S. Freedman: "Testability of Software Components", IEEE Tran. Software Eng. Vol17. No. 6, pp. 553-564. (June 1991).
 [10] G. v. Bochmann, R. Dssouli, J. Zhao: "Trace Analysis for Conformance and Arbitration Testing", IEEE Trans. Software Eng., Vol.15, No.11, pp.1347-1356, (NOVEMBER 1989).
 [11] ISO/IEC JTC1/SC 6/WG 4: "TRANSPORT TEST MANAGEMENT PROTOCOL SPECIFICATION" N507, (1989-04-26).
 [12] PT19@d018f003: "T2CONSCS", ITEX-DE2.1, (Jan. 14 1994).