

トランスポート適合性試験スイート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

連絡先: 〒251 神奈川県藤沢市片瀬山3-11-1 Tel 0466-23-4832

OSIの普及のために、その適合性試験の試験スイートの質の評価と比較が必要であるが、ここで試験スイートの複雑度についての評価を試みる。試験項目の複雑度を試験の実施しにくさと定義し、まずISO 9646で定義された4つの抽象試験法を参考にして試験法のレベルを考える。これによって、よりレベルの低い試験法で実施できる試験目的はより単純であり、レベルの高い試験法でしか実施できない試験目的はより複雑であるとする。この尺度によって分類した試験項目を、さらにシーケンスの長さで細分する。この方法をISOで開発したトランスポートの試験スイートに適用してその評価の妥当性を考察する。

On Estimation of Complexity level of Conformance Test Suites of Transport

Tadao WAKASUGI

WAKASUGI Information Technology Consultant Office

3-11-1 Kataseyama Fujisawa-city, 251 JAPAN

In quality of OSI conformance test suites, there are two categories; necessity and complexity. Here complexity is focused. The levels of complexity of test purposes are classified depending on the levels of testing methods applied to the test purposes. A test purpose is difficult when only high level test methods can use it. Farther each group of test purposes is classified depending on the length of sequence of transitions of the test purpose. By applying the idea to estimating the transport test suites of ISO, the levels show useful as a metric of complexity of test purposes.

1. はじめに

OSI適合性試験の問題の一つに試験スイートの質の評価がある。すなわち、ある適合性試験スイートについて、それが適合性の評価手段として充分か、偏りや足りないところがないか、また他の試験センタで行われている試験の質と同程度か、というような問題である。また適合性試験の実施者にとっては、質のよい試験項目を選択して効率化をはかりたいという要望がある。

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。特にフォールトカバレッジに関するものが多い[1][2]。またIEEEの標準化推進団体である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することを提案している[3]。この試験レベルとComplexityレベルという二つの評価基準は、試験の効果とその実施の困難さを表したもので、試験スイートの評価尺度として基本的なものと思われる。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図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]。

| | High ← Necessity → Low |
|--------------|------------------------|
| Low | Mandatory |
| ↑ | Recommend |
| Complexity ↓ | Optional |
| High | Not recommend |

図1 試験項目の複雑度、必要度と優先度
Fig.1 Copmlexity, necessity and Priority of test cases

本論文では、試験目的の複雑度を試験のしにくさと考え、試験手法とシーケンスの長さを組み合わせて定義し、それにしたがってISOで開発したトランスポートの試験項目の評価を試みる。

2. 試験目的の複雑度

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、状態遷移図における試験項目のシーケンスの長さによるものがある[5][6]。

しかしシーケンスの長さは複雑さの一面しか表していない。シーケンスが短くても合否判定が困難な試験項目がある。またシーケンスの長さは試験手法によっても要求する判定の正確さによっても変わり、意外に客観性に乏しい。

ここでは、複雑度を試験の実施しにくさの程度であるとする。すなわち、試験項目の作成の難しさ、オペレーションの困難さ、試験結果の判定の難しさなども含める。TTCNで記述した試験項目だけではなく、どのような試験法に基づいてTTCNに記述するか、結果判定の記述をどうするかという問題も含め、TTCNで記述できず実施不可能な試験項目も考察の対象とする。したがって複雑度の評価対象は試験項目ではなく試験目的とする。

試験目的と試験項目の関係については、ISO 9646(7)にしたがえば、試験をするにはまずプロトコルに記述された要件のすべてについて試験目的を作成し、それから試験項目を導出する。また試験目的は自然言語で記述し、試験項目は試験項目記述言語であるTTCNで記述することを推奨している。しかし試験目的と試験項目は一対一に対応するわけではない。たとえば、[データ受信待ちの状態にあるIUTが異常な入力に対して正しく応答することを確認する]というような試験目的に対しては試験項目がいくつか作成されるであろう。

IEEEの適合性試験基準[3]では、試験目的または試験項目にあたる概念をアサーション(Assertion)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

この拡張アサーションとは、たとえば、リライアブルな試験方法がないとか、試験の作成や実施に非現実的な量の試験開発時間や手間が必要なものとしている。

3. 試験法と複雑度

試験目的の複雑度はその試験のしにくさの程度であるとする。それは試験法によってレベル分けできる。すなわち、いくつかある試験法をその能力によってレベル付けし、低いレベルの試験法で処理できる試験目的は単純な試験目的で、より高いレベルの試験法でしか処理できない試験目的はより複雑な試験目的であると定義すれば試験目的のレベルが定まる。

試験法のレベルとしては、ISO 9646(7)で定義している、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基準とする。

- (1)R法：R法は下位のPCO(制御観測点)だけで上位PCOのない試験法である。
- (2)D法：D法は、IUTの上下にPCOがある試験法である。上位のPCOがあるだけR法より合否判定の質などが優れている。
- (3)C法：C法は下位PCOとUT(上位テスト)があり、規格化されたTMP(Test Management Protocol)によって規定され制御される試験である。上位PCOについては持つことを義務づけられていないが、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[8]を見ると上位PCOがあるのと変わりはなく、TMPがあるだけD法よりも判定精度や試験実施しやすさで優れている。
- (4)L法：L法は、9646の1994年版では、上位テストと下位テストはIUTとは別システムにある。一方1989年版の9646ではIUT、UT、LTとも同じシステム内にあり、IUTはサービスプロバイダを経由しないでLTとつながっている。これならばタイムの動く時間も正確に測定できる。

以上の検討をもとに、試験手法としての機能の大きいことを記号<で表現すると、次のように記述できる。

$$R < D < C < L$$

ここで試験目的の複雑度を、まず大分類として[あるレベルの試験法では試験できるが、より下位のレベルの試験法では試験できないもの]として分類する。その試験法のレベルとしては、下記の分類を使用する。

- (1)レベル1：IUTの下位のPCOのみを使用して試験を実施し判定できる試験目的。これはR法で実施できる試験項目を想定している。
- (2)レベル2：IUTの下位と上位のPCOとを使えば実施し判定できる試験目的。これはD法を想定したものである。

- (3)レベル3：上下のPCOとそれらを結ぶ試験協調プロトコルがあれば実施できる試験目的。ただし協調作業は下位層のサービスプロバイダを経由するので時間遅れがある。これはC法を想定したものである。
- (4)レベル4：IUPと同じシステム内に上下のPCOとそれらをつなぐテストとがあり、時間遅れのない上下PCOで実施できる試験目的。これは、9646の1989版のL法を想定したものである。
- (5)レベル5：レベル4でも実施できない試験目的、たとえばTTCNで記述できない試験目的などをまとめてレベル5とする。

試験目的の試験法レベルが1～5と上がるにつれ高度な試験機能が必要になることは明かであろう。ここでは、こうして分類した試験目的をさらにシーケンスの長さで分類し、試験法レベルNでシーケンスの長さMの試験目的を、N-Mと記述することにする。

4. 実際分類条件

ここでは前章で定めた複雑度にしたがって実際に分類を試みる。対象はISOで作成したトランスポート試験スイートクラス0の試験目的の英文による記述と、それから導出された試験項目のTTCN記述(8)である。これはC法によって作成されており、TMPのシーケンスも含まれているので読み換えが必要である。したがって、分類に先立ちその読みかえの条件を下記のように定める。以降の文章では主語となるIUTを省略している。すなわち、[受信・・・]は[IUTが受信・・・]ということの意味する。

(1) 試験法レベルのアサイン

同じ試験目的でも合否判定の確実さをどの程度にするかなどによって試験法レベルもシーケンスの長さも異なることがある。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ならばレベル1で実施できるが、上位PCOでの結果もチェックする必要があるとするならばレベル2となり、さらに上位と下位の観測結果の順序も確認すべきであると考えればレベル3の試験目的となる[9]。

ここでは試験目的がいくつかのレベルに入る場合にはもっとも低いレベルをアサインする。たとえば、確立要求に対するIUTの応答を試験する場合、下位PCOからCRを送信し上位PCOでIUTの出力を確認すれば2-1となるが、IUTがCCを送信するのを下位PCOで確認してパスとすれば1-2で実施で

きる。したがってこれは1-2の試験目的に分類する。(2) 試験シーケンスのスタート

試験シーケンスの始まりをどこにするか、アイドル状態からか、テストボディだけにするかなどによって試験シーケンスの長さが変わる。

[5]ではアイドル状態からアイドル状態までの試験実施シーケンス全体の長さを考えているが、そうするとデータ送信や解放試験はコネクションの確立試験よりも複雑ということになる。コネクションを解放するにはまずコネクションを確立しなくてはならないからである。しかしこの評価はあまり適当ではないであろう。

トランスポートは、確立フェーズとデータ転送フェーズと解放フェーズの3フェーズからなるが、ここでは各フェーズの最初をスタートとする。たとえば、[128オクテットで確立したIUTが、オープン状態で・・・]という試験目的は、確立するのに要したシーケンスは数えず、オープン状態から数える。

(3) シーケンスの終了

シーケンスのスタートと同様に終わりをどこまでとするかが問題である。ISOの試験スイートでは、切断のテストでは切断の確認のためのシーケンス、確立のテストでは確立の確認のためのシーケンスが付加されるので、実際に試験対象になるシーケンスよりも長い。さらにパスの後にアイドルに戻るシーケンスがつく。

本論文ではTTCN記述でPASSという判定記述のあるところまでとする。すなわち、確認のためのシーケンスは含め、それ以降のシーケンスは含めない。

(4) TCP (試験協調手続き) の扱い

ISOの試験スイートはC法で記述されているので、TCPの手順が含まれる。しかしこれらはシーケンスの数には入れないことにする。

(5) シーケンスが一意に定まらないケース

一つの試験目的に対しいくつものシーケンスが考えられる場合は、最も短いものを採用する。

たとえば[IUTが、オープン状態で異常なCR__TPDUを受信して、ER__TPDUを送信するか、ネットワークを切断するか、または無視することを確認する]という試験目的の場合、[切断する]の判定には切断情報を受信するという1シーケンスの確認がすむが、[無視する]の判定にはLTからデータを送信して応答があるかで判定するので2シーケンスが必要である。したがって切断の方を採用する。

またデータの分割送信や交互送受信の試験などは試

験項目作成者によってシーケンスの長さが異なる。これらもISOで記述したものの中からもっとも短いシーケンスのものを採用する。

5. トランスポート試験目的の分類

ここで具体的にISOで開発したトランスポートクラス0[8]の試験スイートをレベルに分ける。ISOで開発した試験スイートは次の4種114件からなる。

- (1)有効な機能試験 (Valid capability test) : PICS (プロトコル実装適合性記述書) に、実装したと書かれている機能が実際に存在するかどうかを確認する試験
- (2)有効な振舞い試験 (Valid behaviour test) : 動的適合性要件をどの程度まで満たしているかを確認する試験
- (3)無効試験 (Invalid Test) : 規定の少なくとも一つに違反しているイベントに対する確認試験
- (4)インオポチューン試験 (Inopportune Test) : あるイベントを、規定によって発生が認められていない時点で発生させて結果を確認する試験

ISOで開発したトランスポートクラス0の試験項目の分析では、試験法レベル2までしかない。したがってまず有効試験についてレベル2まで考察し、そのあとで無効試験とインオポチューン試験について考察し、最後にレベル3以上について考察する。

5.1 有効試験の複雑度

(1)レベル1

このレベルは下位PCOしかないので、LT側からコネクション確立要求をしてIUTの応答機能を試験することはできるが、IUTの確立要求機能は試験できない。IUTの応答試験についてもパラメタの指示や制御などはできない。

・1-2

LTからCRを受けてCCを返したり、またはDR/ERを受信してなんらかの反応をするというような二つで一組のシーケンスについて、いろいろとパラメタを取り替えて試験するものが18件ある。

・1-3

ここには4件の試験目的が属するが、内容は正常解放やTS2タイマ機能などの切断試験である。IUTがネットワークを解放したかを確認するためにLTが確立要求するというシーケンスを使うので長さが3と

なる。

(2)レベル2

このレベルでは上下のPCOが使用できるのでIUT側からの起動機能が試験できる。

・2-1

IUTが起動側となる単純な確立/切断要求とデータの送受信の試験が7件ある。

・2-2

レベル1-2と同じように2つのシーケンスで1組となる試験目的で、かつ試験の実施や判定に上位PCOでの指示や観測データが必要な試験目的10件がある。ネゴシエーション機能試験が大半である。その他に、タイマTS1、上位ユーザからの確立拒否、データシーケンスの機能試験などである。

・2-3

レベル2-2よりもやや複雑なもの10件がここに入る。うち5件は、LTが指定した長さのデータをIUTが送信するかを確認する試験で、確認のシーケンスのために長さが3となる。残り5件はWFCC状態での機能の試験で、WFCCの状態にまでもってゆくシーケンスのために長さが3となったものである。

・2-4以上

レベル2-4以上は数が少なく、反復性のある試験目的と分割/粗立の試験目的がいくつかある。

レベル2-4の1件は確立と解放の反復試験である。レベル2-5には、分割/粗立の試験でシーケンスの数が多くなったものが2件と、確立のやり直しをするという試験が1件ある。

トランスポートのクラス2ならば、連結/分離、多重化/逆多重化、結合/分解などの試験目的がある。

5.2 無効試験とインオポチューン試験の場合

無効試験目的は有効試験に比べると単純で、次に示すように7種類に分類される。CR/CC/DTの各TPDUの無効パラメタに対する試験目的と、クローズド/WFTRESP/WFCC/オープンの各状態でのインオポチューン試験目的である。

・1-2

①CR-TPDUの無効パラメタ試験が10件

②DT-TPDUの無効パラメタが7件

③オープン状態でインオポチューンなTPDUを受信するものが8件。

・1-3

④クローズド状態でインオポチューンなTPDUを

受けて処理するものが7件、IUTがそれを無視したかどうかの確認のために、LTは不正なTPDUの送信後に正常なCR-TPDUを送信してCC-TPDUが送られてくるのを待つので、長さ3になる。

- ⑤WFTRESP状態でのインオポチューン試験が10件、IUTをWFTRESPの状態にもって行くためにシーケンスの長さが増えた。
- ・2-3
- ⑥CC-TPDUの不正パラメータ試験が12件
- ⑦WFCCでのインオポチューン試験が7件

インオポチューン試験では、IUTの状態によって複雑度が異なり、

オープン(1-2)

<クローズド(1-3)=WFTRESP(1-3)

<WFCC(2-3)

となる。括弧内は複雑度レベルを表す。

オープンよりもクローズドの方が複雑なのは、試験実施後の状態を確認するために確立要求をするからである。WFCCがさらに複雑なのは、WFCC状態にもって行くためにCRの受信が必要なためである。

5.3 レベル3以上の試験目的

ISOのトランスポートクラス0の試験スイートにはレベル3以上の試験目的はない。しかしクラス2などでは次のような試験目的が考えられる。

(1) レベル3

上位下位PCOでの情報の順序関係を確認する必要がある試験目的がここに属する。クラス2の明示的フロー制御(ウィンドウ制御)やExpedited Dataと通常データの混合送受信などがある。

(2) レベル4

外部要因、特にサービスプロバイダの影響を受け実施/判定が困難な試験目的はこのレベルで試験可能となる。

タイマの試験がここに入る。作動時間が予定どおりかのチェックまでするならばレベル4の試験手法が必要である。

(3) レベル5

いままで述べた試験目的の他に、適合性試験では実施しないか、IS9646のルールで適合性試験の範囲外とされているが実際に実装製品を開発する場合には必要という次のような試験目的がある。

- ・実装規約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ことを確認する試験。たとえば実装規約でTPDU長256オクテットと決めても、相手側が512を希望しIUTも512を受けることができるならば512と決めることができ、それをエラーとはしない。
- ・IUTが起動側の場合、実装規約でノットアプリケーションと決まっても、上位テストでその使用を抑制できない機能を、IUTが使用しても不合格としない。
- ・IS9646では、性能、信頼性、ロバストネス(Robustness: 頑強さ)試験、リアルイフェク(Real Effect)の試験などであるとしている。ここでロバストネス試験とは、わざと間違った入力をして反応を見るというような意地悪試験で、上位層から不正要求をしてIUTの反応を見るというような試験がここに入るであろう。

6. 分類結果の評価

トランスポートクラス0についての上記の分類結果を表1に示す。

表1によると、試験目的のほぼ半分づつを複雑度レベル1と2でカバーしている。有効試験ではレベル2が多く、無効/インオポチューン試験ではレベル1が多いが、特にどこに集中しているということはない。

このレベル2までに含まれない試験目的は、タイムの時間測定がある程度である。したがって、C法やL法という高度な試験は、試験の範囲を拡大するためにはあまり必要ではなく、むしろ判定の正確さや試験のしやすさという点にその主目的があると言える。

またこの表ではシーケンスの長さは5以下である。これはISOの試験は機能を一通り試験しているものであって、各種機能を組合せた試験や、機能の反復による信頼性試験などはあまり含まれていないためである。

7. まとめと今後の課題

本論文では、OSI適合性試験目的の複雑さのレベルを試験手法とシーケンスの数で分類し評価することを試み、その方法をトランスポートの試験スイートに適用した。

この方法は、シーケンスの長さだけで複雑度を評価する方法よりもきめ細かい評価となり、実感に近いものが得られたと考える。また表1から、トランスポー

| 複雑度 | 試験レベルの内容 | Valid Test Cases | 件数 | Invalid & Inopportune Test Cases | 件数 | Total |
|-------|-----------------|----------------------------|----|----------------------------------|----|-------|
| 1-2 | 下位PCOのみで試験可能なもの | 確立の受諾/拒否, DR/ER処理など | 18 | 無効CR/DT処理, オープン状態のインパクション | 25 | 64 |
| 1-3 | | 切断, TS2タイマ試験など | 4 | WFTRESP/クロズドのインパクション | 17 | |
| 2-1 | 上下のPCOで試験可能なもの | 確立/切断要求, データ送受信など | 7 | 無効CC処理, WFCCでのインパクション | 19 | 50 |
| 2-2 | | ネゴ機能, TS1タイマ, データシーケンス機能など | 10 | | | |
| 2-3 | | DT-TPDU長の指定, WFCCでの機能など | 10 | | | |
| 2-5 | | 分割/組立, 確立のやり直しなど | 3 | | | |
| 3 | TMP使用 | 明示的フロー制御など | - | | - | - |
| 4 | 同一システム内 | タイマの作動時間確認など | - | | - | - |
| 5 | 試験不可 | 信頼性試験, IUTの内部状態確認など | - | | - | - |
| Total | | | 53 | | 61 | 114 |

表1. 複雑度と試験項目数(トランスポート試験スイートクラス0の場合)

Table 1. Complexity levels and no. of test cases (A case of Transport class 0 test suite)

トランスポートの場合の試験目的の複雑度による分布状態が見え、試験スイートの構造の一面が示せたと考える。今後の課題としては、また試験目的の必要度による評価と組合せて、試験目的の質の評価体系の確立をほかりたいと考えている。

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.119-32, (Oct. 1991).
- (2) R.Castanet, P.Corvisier, R.Casadessus: "Industrial 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, pp.119-129, (Oct. 1990).
- (5) A.Petrenko, R.Dssouli, H.Koenig: "On Evaluation of Testability of Protocol Structures", Proc. of IWPTS#6, pp. 114-127, (Sep. 1993).
- (6) R.S.Freedman: "Testability of Software Components", IEEE Tran. Software Eng. Vol17. No.6, pp.553-564. (June 1991).
- (7) 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).
- (8) PT19@d018f003: "TOCONSCS", ITEX-DE2.1, (Jan. 14 1994).
- (9) 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).