

OSI 適合性試験スイートの評価法 — マルチトランジションカバレッジ —

若杉忠男

技術コンサルタント

OSI 適合性試験の試験スイートの比較評価の方法の一つとして、ここにマルチトランジションカバレッジ法を提案する。この方法は、プロトコルをFSMで表現し、その連続するn個のトランジション単位で合否判定をするものである。nが多いほど試験の質が高いと考えられ、その程度によって試験スイートの質の評価ができる。またこれを試験スイート作成のガイドラインと考えることもできる。

1. はじめに

OSIの適合性試験の試験スイートについては、国際標準であるISO9646にその開発方法の指針が定められている[1]。それにしたがって、多くの試験スイートが開発されてきた。一方、通信プロトコルの適合性試験手法として、通信プロトコルをFSM (Finite states machine : 有限状態機械) で表現し、その各トランジションを実行するように入力データ列を作成してIUT (試験対象) に適用するという方法がよく研究されている。T法、UIO法、D法、W法などがそれである[2][3]。

また試験スイートについて、その国際協調という見地から、質を比較し評価する必要性は以前から指摘され、各種のカバレッジやメトリックが提案されている[4][5]。本論文で新たに紹介するマルチトランジションカバレッジ法はその評価法の一つであるが、また試験項目作成基準の一つともなりうる。

一般にすべてのフォールトを発見できる方法は有り得ないが、ここで提案するマルチトランジションカバレッジ法にはレベルと

いう評価尺度があり、このレベルを上げ下げすれば、試験に要する労力と発見できるフォールトとの関係が分かるという利点がある。

以下、まず2節でFSMのカバレッジ、3節でマルチトランジション試験の考え方を紹介し、4節で特にレベル2のマルチトランジションカバレッジについての例を示す。5節では本方法の基本的な考え方を述べ、6節で各レベルと発見できるフォールトとの関係を説明し、7節で試験項目を具体的に示し、8節ではトランジションによるトランスポートの試験スイートの質の評価例を示し、最後に9節でまとめと今後の課題を述べる。

2. FSMのカバレッジ

一般のプログラムの試験には、プログラムカバレッジという方法がよく用いられる。これは、プログラムをフローチャートで表現し、その各ステートメントをすべて通過するような入力データを作成して試験するものである[7]。通信プロトコルについて

も、これと同じ様な考えで試験項目を作成することができる。ただしここでの対象はフローチャートではなくプロトコルを表現したFSMである。プロトコルを表現するのに使われるFSMとは、図1に示すように、状態(state)を表す丸、そこからのトランジション(transition)を表す矢印、その矢印に附加した入力/出力データ、トランジション先の状態(end state)からなるものである。

図1は、IUTが状態1にあるときに、入力1を与えると出力1を出して状態2に遷移し、さらに入力2を施すと出力2を出して状態3に遷移するということを表している。トランスポートプロトコルの例でいうと、アイドル状態のIUTにトランスポートCRリクエストを与えると、CR-TPDUを送信してWFCC(CC待ち)状態になるというようなものである[6]。

国際的に定められたOSIの適合性試験は、ブラックボックス試験である。したがって、図1で状態1にあるIUTに入力1を与え、出力1を観測することはできるが、遷移した先が状態2であるかどうかは簡単には確認できない。したがってそれを確認するために、状態2にあると思われるIUTに引続いていくつかの入力を与え、観測される出力がプロトコルで定められた状態2特有のものであれば状態2と認めるという方法をとる。しかしこれは、後でも述べるが、さまざまな条件が満たされないと成り立たない。

このFSMのカバレジについては、次の

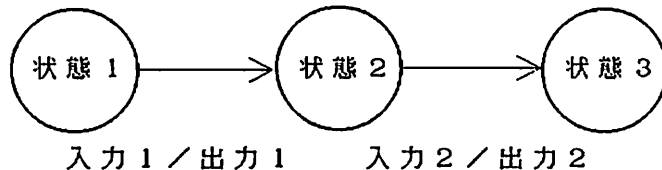


図1 FSMのトランジションと状態

ようないくつかの考え方がある。

- ・ステートカバレジ：これはすべての状態を一度は通るような試験項目を作成するもので、MT法 (Modified Transition Tour) とも呼ばれている[8]。これは試験したい状態にコントロールをもって行くためにも使われる。
- ・トランジションカバレジ：これはすべてのトランジションを一度は実行するように試験項目を作成するものであり、T (Transition Tour) 法と呼ばれる。
- ・エンドステート確認付きトランジションカバレジ：適合性試験としては、トランジションカバレジを満たすだけでは不十分で、エンドステートが予定のものかどうかを確認する必要がある。その確認のための試験項目をトランジションの試験項目の後に追加する。UIO法、W法、D法などがこれである。

3. マルチレベルトランジション試験

ここでは、一つのトランジションについてその動作の正しさを確認する試験をレベル1のトランジション試験と呼ぶことにする。それに対し、2つの連続したトランジションを組合せてその結果を評価することをレベル2のトランジション試験、またいくつかのつながったトランジションをまとめて評価する方法を、マルチレベルトランジション試験と呼ぶことにする。

一般にマルチレベルトランジション試験は次のような場合に使用される。

(1) 複雑な条件の試験

試験目的には、もともとトランジション一つでは試験できないものがある。いくつかの動作の組合せや、反復動作の試験などである。以下、トランスポートの例で説明する。

- ・いくつかの条件の組合せ試験：ウィンドウ制御の試験など
- ・大量データの連続処理試験：大きなサイズのTSDUをいくつかに分割して送信するセグメンテーションなど
- ・反復した動作の試験：普通データと優先データを、交互にあるいは相互に送受信するなど
- ・その他：コネクションの多重化など

(2) 間接判定

試験システムによっては、レベル1のトランジション試験では判定に必要な出力が得られないことがある。たとえば、トランスポートをR法で試験する場合、IUTにコネクション確立要求を出してそれが受理されたかどうかの確認は上位PCOがないので簡単にはできない。その場合には、次のトランジションであるCC送信を待つ。IUTからCC確認が来れば、コネクション確立要求をちゃんと処理していると判断する。

ここでレベルIの間接判定という言葉を定義する。

- ・レベルIの間接判定：レベルIの判定で、1、2、3、・・・Iの出力のうち、いくつかを省略した残りが予期のものと等しければ合格とする判定法。

例えば、トランスポートコネクションの確立を、下位テストからCRを送信してCCが送信されてくるかどうかで判定するの

は、レベル2の間接判定である。また、出力がないのが正しい結果だというような試験項目がある。たとえば、エラーのあるパラメタを含んだTPDUを受信したIUTがそれを無視するという試験項目である。そういう場合には、状態の変化が起きていないかの確認をするトランジションを付加した試験をする。

一連のトランジションの最後の出力だけを観測して、それが正しいければ経路すべてのトランジションが正しいとする判定もある。これは一般のシステム試験などで行われている考えで、ブラックボックステストの中のブラック度の高いものと見ることができる。

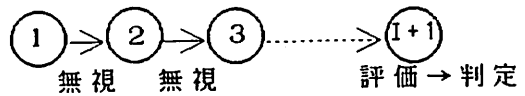


図2 レベルIの間接判定（中間を無視）

4. レベル2トランジションカバレジ

ここでは、特にレベル2トランジション試験について考える。まず次の定義をする。

- ・レベル2のトランジションカバレジ：2つの連続したトランジションの組み合わせすべてについて、レベル2のトランジション試験をすること。

この試験の狙いと必要性は次のとおりである。

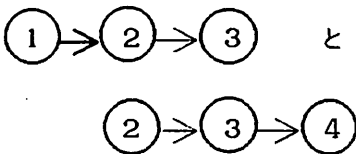
(1) エンドステートの確認

先にも述べたように、ブラックボックステストでは、1トランジションの試験で図1の状態1から入力1で到達したエンドステートが状態2であるという確認が難しい。確認のために、状態2に達した後でさらに入力2を入れてみる。これで出力2でないものが出たらフォールトである。出力2が

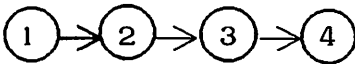
得られたら状態2が目的の状態である確率は高い。しかしそれでも100%確実ではない。

確実性を高めるには、さらに入力が続けるということが考えられる。トランジションの数を増やせば手間が増えるだけ判定精度はよくなる。すなわち、状態1→2→3と試験するだけでは確実でないというならば、1→2→3→4と試験すれば、それだけ確実性は増す。UIO法、D法、W法などの考え方も、これと同じようにエンドステート確認のためにいくつかの入力を実施しているが、これらの方法では、完全定義とか、IUTのFSMの数に一定の制限をもうけて、その範囲内で100%のカバレッジを保証する。

レベル2のトランジションカバレッジの考え方は、2つ並んだトランジションの組合せすべてを試験すれば、例えば、1→2→3と2→3→4の2ケースのレベル2の試験をすれば、1→2→3→4と連続して試験したものと実質的に同じになるだろうと考える。



の試験は



と同じ効果があるとする。

図3 レベル2のトランジションカバレッジの考え方(1)。ここで確認する対象は、状態1と状態2の間のトランジションである。

すなわち、「1→2、2→3、3→4という3つのレベル1試験をパスすれば、1→2→3→4もパスするだろう」というのは、少し論理に飛躍があるが、「1→2→3と2→3→4の2ケースのレベル2の試験をパスすれば、これらの試験項目は、2→3のところと同じことを試験しているのだから、1→2→3→4もパスするだろう」と考えるのは受け入れやすい。

こうして、すべてのトランジションが正しいといえれば、IUT全体が正しいと言える。

(2) 状態単位での正しさの確認

(1)で述べた試験の目的は、状態1から状態2に行くトランジションの正しさの確認である。すなわち試験はトランジション単位に考えられ実施されている。それに対し、見方を変えて、状態単位での正しさの確認をすることを考える。たとえば図1で、レベル1の二つの試験項目をコンカチネートして作った試験項目、「入力1」・「入力2」を実施し、「出力1」・「出力2」となることを確認すれば、これらのトランジションの間にある状態2が正しいと判定する。入力トランジションが2個、出力トランジションが3個あれば、2×3件の試験をすることになる。

IUTが、状態別にモジュール化されている場合には、状態単位の試験というアプローチは、すなおな考え方である。こうして、すべての状態が正しいと確認できれば、IUT全体が正しいとする。

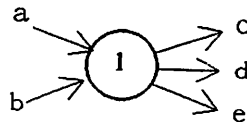


図4 レベル2のトランジションカバレッジの考え方(2) (確認する対象は状態1の動作である)

5. マルチトランジションカバレッジ法の考え方

ここでは、一般化してマルチレベルのカバレッジについて考察するが、それに先立ちまず言葉の意味を説明する。

- ・レベルIのトランジション試験項目：引き続きI個のトランジションを試験する試験項目のこと。たとえば、「 J 」を試験項目、「 J 」・「 J 」を試験項目のコンカネーションを表すとすると、2レベルのトランジション試験とは、「コネクション確立要求」・「CC受信」とか、「通常データ送信」・「エクスピダイトデータ送信」とかである。
- ・レベルIトランジションカバレッジ：レベルIトランジション試験項目で、FSM全体をカバーする試験または試験項目を表し、記号M10で表す。I=1のとき、これは通常言われる、トランジションカバレッジとなる。
- ・異常入力試験項目：状態に異常な入力を与えて、プロトコルに決められていないようなことが起きないことを確認するための試験項目。これはインオポチューンテストイベント（起きることが仕様で許されていない時に起きるイベント：コネクションを確立していないのにデータを送信してくるというようなもの）とインバリッドテストイベント（少なくとも一つの適合性要件を犯しているイベントで、パラメタの誤りとかシンタックスエラーなど）、その他不都合を起こすものを含む。
- ・レベルJの異常入力試験項目：J-1個の正常なトランジションの後で異常入力試験をする試験項目。遷移してきた経路によって異常入力に対する処理が変わらないことを確認し、異なった経路を経て到達しても、異常な入力に対して正しく

処理することを確認するために使用する。

- ・レベルJ異常入力カバレッジ：レベルJの異常入力試験項目で、FSMをカバーする試験または試験項目を表す。記号M0Jで表す。

ここで、記述法を統一するために、M00はステートカバレッジ試験項目を表すこととする。またM10とM0Jの試験項目を組み合わせた試験項目をM1Jと表わすことにする。

さて、一般に、試験スイートの質は、試験項目の数だけでなくその実施順序や組合せ方によっても変わる。たとえば、いくつかの試験項目をコンカネートして作成した試験項目は、もとの別々の試験項目のグループよりも、連結しても実行できるという確認があるので、それだけ情報量が多く試験項目として質がよいと考えられる。XがYより質がよいことを $Y < X$ と表し、試験項目のグループを()で表現すれば、次のように書ける。

$$(\text{「A」}, \text{「B」}) < (\text{「A」} \cdot \text{「B」})$$

そうすると、マルチトランジションカバレッジはレベルが上がるほど試験項目が増え、その質が上がるから次のように記述できる。

$$M0J < M1J < M2J < M3J \dots$$

$$M10 < M11 < M12 < M13 \dots$$

一般に、 $I < K$ で $J < L$ ならば、 $M1J < MKL$ である。

すなわち、マルチトランジションカバレッジ法では、レベルが増加するにしたがって試験項目の質が上がり発見出来るフォールトが増え、その代わり試験の手間も増加する。これによって階層的な試験項目体系が作られ、試験項目の質の分類ができる。したがってこれを使って試験スイートの質を評価し、「この試験スイートは、M10の試験項目がx件」とか、「M20のカバレッジ

Y%程度」などと、レベルを件数またはカバー率で表現できる。

6. 各レベルとフォールトの関係

前節で述べたレベルと、発見できるフォールトとの関係を、図5のFSMを参照しながら説明する。この図5の実線部分をプロトコルで定められたものとし、いま状態3が正しいことを確認したいとする。

- (1) M00はスタートカバレジで、すべての状態の存在とそこへの到達経路を確認できる。
- (2) M10では、すべてのトランジションの存在と動作の確認ができる。「入力1」「入力2」「入力3」を適用し「出力1」「出力2」「出力3」を確認する。しかしこれではエンドステート確認がない。
- (3) M20の試験項目により、各状態が正しい入力には正しく動作し、かつエンドステートについてもそうとうの確実さで確認ができると考えられる。しかし異常入力があったときにどうなるかについては、なにも確認していない。
- (4) M01によって、各状態に異常入力をして不都合なトランジションが起きないことを確認できる。

図5の入力4がそれである。正常入力に対して正常動作をするだけでは、異常な入力があった場合に何が起きるか分からない。異常入力としては事故や天変地異を含め何が来るか分からないので、いろいろなケースで試験しなくてはならず、件数が多くなりがちである。従来、正常入力の試験項目について、件数を減らす工夫はよく検討されているが、この異常入力試験項目の件数も無視できない。状態3に外部から異常に飛び込んでくるというフォールトもあるが、これは飛び込んで来ても事故を起こさなければよいから、出力トランジションのチェックをすれば入力トランジションのチェックは必要ない。

このM01とM10を組み合わせたM11で、FSMの一応の確認はできる。

M10よりレベルが一つ上のM20では、完全とはいえないがエンドステートを含めた確認ができ、さらにM21では、異常入力試験を含めた状態ごとの動作確認試験ができて、ほぼ完全なFSMの試験になると考えられる。

ここで、マルチトランジションカバレジ

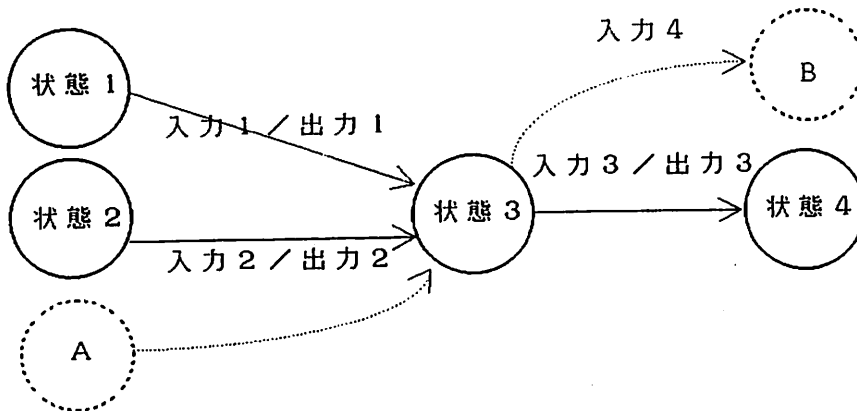


図5 FSM (2レベルの場合の説明)

と、それによって確認できる事項及び確認できない事項を表1に示す。各レベルごとに発見できるフォールトとできないフォールトのタイプを明確にして、それを発見したければもっとレベルを上げて費用をかければよいというのが、マルチトランジションカバレッジの考え方である。

7. 試験項目シナリオと試験実施の手間

マルチトランジションカバレッジ法の試験項目について、実用性を考慮して考察する。トランスポート層の例を用い、M10とM20

で説明する。実用上、この二つが適合性試験項目として適当であろう。

実際に使用する試験項目のシーケンス（これを試験シナリオと呼んでいる）では、初期状態からスタートして初期状態に戻るまで試験し、レベル1の試験ならば1トランジションの単位で合否を考え、レベル2の試験ならば2個の単位で合否を考える。たとえば、「CR要求発行」「CC確認」「データ送信」「切断要求」の4つの要件を試験するのに、レベル1とレベル2では

表1 試験項目のレベルと質

(1) 正常入力の場合

レベル	確認できる事項	確認できない事項
M00	ステートカバレッジ。仕様に定められた状態とそこへの経路の存在の確認	遷移すべてについては確認していない
M10	仕様に定められた遷移が正常な入力に対し正常な出力をすることの確認	余計な遷移があるか、遷移の到達先が予定の状態かは不明
M20	エンドステートについてのある程度の確認を含めすべての遷移の正常動作	3つ以上先の遷移との関係/影響についてはチェックしない
M30	経路のうちの3つ連続した遷移が正常入力に対し正しく出力することの確認	4つ以上先の遷移との関係/影響についてはチェックしない

(2) 異常入力の場合

レベル	確認できる事項	確認できない事項
M01	各状態が異常な入力に対して有害な遷移をもたないことを確認	通ってきた経路により結果が変わるかもしれないということは考慮しない
M02	各状態が遷移してきた経路に関係なく異常入力を同様に処理することの確認	一つ前の経路の影響を確認するが二つ以上前の経路の影響までは確認しない

表2のようになる。

表2 試験シナリオ例

レベル1：「CR発行」・「CC確認」・
「データ送信」・「切断要求」

レベル2：「CR発行」・「CC確認」・
「データ送信」・「データ送信」・
「切断要求」・「CR発行」

レベル2では、判定は隣接する二つのトランジションが合格ならば、その間に入った状態を合格とすると考える。たとえば「CR発行」・「CC確認」が合格ならば「CC待ち状態」が合格とする。表2で、レベル1と違うところを説明すると次の2箇所である。

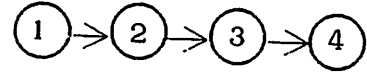
一つは「データ送信」・「データ送信」と2回繰り返す、2度連続送信可能なことを確認し、「オープン状態」が正しいとする。また最後に「切断要求」・「CR発行」として、アイドルに戻ったことを確認して「アイドル状態」の判定をする。

この例では、レベル1の試験項目ではイベントが4つで、レベル2の試験項目では6個となった。それだけレベル2の方が項目数が増加する。しかし、その増加件数はたいしたものではない。すなわち、図6の(1)のような場合には手間の増加はわずかである。

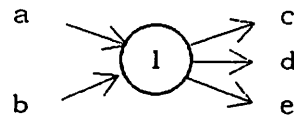
一方、図6の(2)のように一つの状態にn個の入力とm個の出力がある場合には、M10ではn+m個のトランジション(a、b、c、d、eの5ケース)を試験すればよいが、M20では入出力すべての組合せでn×m個(a・c、a・d、a・e、b・c、b・d、b・eの6ケース)の試験をしなければならない。

トランスポートクラス0では、一番入出

力の多い状態は「データ転送待ち状態」で、入力トランジションは2個(接続応答、接続確認)、出力トランジションは3個(切断要求、切断指示、N-リセット指示)で、さらにこの状態から出てまた戻るトランジションが2個(データ転送要求、転送指示)ある。



(1) 1入力1出力の場合



(2) 多入力多出力の場合

図6 トランジションの型

それらの組合せを全部実行すると4入力5出力で、試験項目数はM10で9件、M20で20件になる。

これにさらに先に述べたようにデータ転送を2度ずつ実施するとか、データ送受信を交互に実施するとかすれば、さらに増加する。すなわち、試験に要する労力は、M10とM20とでは2倍ていどの増加となる。

8. 試験スイートの質の比較

ここでは、このトランジションのレベルを、既存の試験スイートの質の評価指標に使用することを考える。試験スイートの質の比較指標として先にも述べたように各種のカバレッジが提案されているが、既存の試験スイートを簡単に評価できる指標は、試験項目の件数くらいしかない。

トランジションのレベル別の件数とその試験項目に占める割合を調べれば、複雑な条件の組合せをどの程度ていねいに試験し

ているかが評価できるであろう。その手順は次のとおりである。

- (1) 評価対象を、正常試験項目に限定する。不正入力試験やインオポチューン試験は、トランジションレベルを上げる意味が少ないので考慮外とする。また試験方法はD法によるものとする。R法による試験スイートとD法によるものをそのまま比較はできない。
- (2) レベル1、2、および3以上のクラスに分けて、その件数と割合をカウントする。
- (3) 高レベルの試験項目件数が多い方が、複雑な条件の試験をていねいに実施しているといえる。また、レベル2の試験項目については、間にはさまった状態別に分類すれば、どの状態の試験項目の質が少ないかが分かる。

表3 トランスポート試験スイートの比較例（ISOの試験スイート）

クラス	クラス0		クラス2	
	件	%	件	%
Ⅷ#1	37	69.8	103	66.9
Ⅷ#2	15	28.3	24	15.6
3以上	1	1.9	24	15.6
判定不可	0	0.0	3	1.9
合計	53	100	154	100
状態の数	4		6	
件/状態	13.3		25.7	

実際に、マルチレベルの試験項目がどのくらい用いられているかを見るために、ISOで開発しているトランスポートの試験スイートを調査した[9]。

ISOの適合性試験のガイドライン[1]には、適合要件を一つ一つ記述して、それをもとに試験目的を作成し、それらをもとに試験項目を作成するように推奨している。いくつかの試験目的を一緒にして複合試験項目を作ってもよいとしているが、実際にISOで作成した試験スイートを見ると、大半がレベル1の試験項目である。

このトランジションの数を数えるという作業は、資料自体がまだ完全でないためもあるが意外に難しく、主観的判断に左右されるような点も多いので、いくつかのルールをもうけて分類した。それは次のようなものである。

- (1) 試験はD法によるとした。すなわち、上下のPCOを完全に制御し観測できるとした。
- (2) 試験項目の前提条件は、トランジションの中には数えなかった。たとえば、「128オクテットで確立したIUTが、オープン状態で・・・」という試験項目は、確立するまでに要したトランジションは数えず、オープン状態をスタート状態としてそこからトランジションを数えた。
- (3) いくつかのケースが考えられる場合には、もっとも短いものを採用した。たとえば、「IUTが、クローズの状態、不正なサイズのCRTPDUを受信して、ERTPDUを送信するか、DRTPDUを送信するか、ネットワークを切断するか、またはCRTPDUを無視することを確認する。」という試験項目については、「無視する」を採用した。
- (4) 「・・・を受信して無視する」はトラン

ジション1個とした。

- (5) 正常解放は、クラス2ではレベル2トランジション、クラス0では1レベルとし、また異常解放は両クラスとも1トランジションと数えた。
- (6) 「いくつかのDT T P D Uを・・・」というようなトランジションがいくつになるか不明なものは、すべて3以上に分類した。

表3がその結果をまとめたものであるが、これから次のようなことが分かる。

- ・試験項目件数から見ると、クラス2の方が圧倒的に多い。状態あたり件数でもクラス2が倍近くある。
- ・レベル1が全体に占める割合は、両クラスともほぼ等しく、2/3である。
- ・レベル2と3については、クラス2でのレベル2の割合が少ないが、それ以上にレベル3以上の割合が増えている。

以上の結果から見て、件数ではクラス2がよく、また内容についてもクラス2の方がややよいと言える。これはトランスポートのプロトコルが、クラス2の方がクラス0よりもずっと複雑になっているためでもある。

9. おわりに

マルチトランジションカバレジ法の考え方や試験項目の質と評価について述べたが、いくらカバレジを検討しても、フォールトの100%を発見できるわけではない。現実に実行できない試験項目もある。本マルチトランジションカバレジ法では、どのレベルまで試験するとどのようなフォールトは発見でき、どのようなフォールトは発見できないか、またどのくらい手間が増えるかについてまとめ、検証試験の限界と責任範囲を明確にすることを目指した。

またこの考えを試験スイートの評価に使い、既存の試験スイートの評価ができることをトランスポートの例によって示した。

今後、さらに比較対象を広げ、その実用性を検証する必要があると考えている。

参考文献

- [1] Information technology - Open System Interconnection - Conformance testing methodology and framework- Part 2: Abstract Test Suite Specification, ISO/IEC JTC 1/SC 21, ISO 9646-2, 1994-3-14.
- [2] S.Fujiwara and others, Test Selection Based on Finite State Models, IEEE VOL.17, NO.6, JUNE 1991
- [3] D.Sider and others, Formal Methods for Protocol testing: A Detailed Study, IEEE VOL.15 NO.4, April 1989.
- [4] G.v.Bochmann and others, Faults Models in Testing, IWPTS #4, Oct.15-17 1991.
- [5] S.T.Vuong, J.Curgus, Test Coverage Metrics for Protocols, IWPTS #4, Oct. 15-17 1991
- [6] ISO 8073 Information Processing System - Open Systems Interconnection - Connection Oriented Transport Protocol Specification, 1988-12-15.
- [7] Boris Beizer著、小野間、山浦訳、ソフトウェアテスト技法、日経BP出版センター、1994年2月刊。
- [8] 佐藤、宗森、勝山、水野、通信システムの段階的な試験のための試験系列自動生成手法とその実現、情報処理学会論文誌、Vol.31, No.10, Oct.1990
- [9] EWOS PT19, TP/TSS for ISO 8073 Class 0/2/4 based on ISO 10025-2, ITEX-DE2.1 TOCONSCS, Jan. 14, 1994.