

著者紹介

猪飼 維斗 (学生会員)

2008年東京海洋大学大学院海洋科学研究科修士課程修了。2009年より株式会社エム・アイ・ベンチャー社員。2011年電気通信大学大学院情報理工学研究科入学。道路交通シミュレーション，市場取引の分析などの研究に従事。日本シミュレーション学会，情報処理学会などの会員。

板倉 直明 (非会員)

1985年慶応大学理工学部卒業。1987年慶応大学大学院理工学研究科修士課程修了。1990年慶応大学大学院理工学研究科博士課程修了，博士(工学)。1990年より電気通信大学。ブレインコンピュータインタフェース，生理的振戦，視線入力システム，道路交通シミュレーション，などの研究に従事。計測自動制御学会，日本生体医工学会，電子情報通信学会，日本生理人類学会などの会員。

要旨

今日、多くの使いやすさを目的としたマルチエージェント型シミュレーションが存在している。しかしモデルの作成には思考時間によって生じる作業のばらつきがあり、作業時間から使いやすさを定量的に評価することは難しい。そのため本研究では総作業時間と Operator 作業時間、思考時間に着目しユーザの感じる使いやすさを評価する指標を考えた。

本研究では総作業時間から実際に作業の動作を行った時間である Operator の時間を差し引くことで残余時間を計算した。この残余時間は課題の認識や手法選択の判断の時間を含んだもので、主に思考時間によって構成されている。そして残余時間とモデル作成作業への満足度との比較をおこなった。

本研究ではこの方法を 3 つのシミュレーションシステムに対して適用し定量的な評価を試みた。その結果、3 つのシミュレータ間で総時間では有意な差があったものが、残余時間では有意な差が確認されなかった。残余時間を用いて評価する事で、思考時間が主となる作業においても定量的な評価が可能となる事が示された。

Abstract

There are many Multi-Agent Simulation systems that are intending easy to use. When we will evaluate modeling usability now, it is difficult to evaluate quantitatively from operating time, because there are variability in tasks caused by thinking time on construction of models. So we designed to evaluate usability from “the total time”, “the thinking time” and “the operating time”.

In this study we calculate “the remain time” by subtract operating time from total task executing time. The remain time is including the meta thinking time, that is, decision of the method and cognition of the task goal. And we compared the total time, operating time and the remain time, with a satisfaction rate of modeling.

We applied this method to 3 simulation systems to evaluate quantitative usability. Then there confirmed that a significant difference between 3 systems with the total time, however, there are only non significant differences between 3 systems with the remain time. This study shows that there is correlation between the remain time and the satisfaction rate. And estimated when using the remain time, It can be possible to evaluate quantitatively about tasks with much thinking time.

1. はじめに

シミュレーションによる解析や結果の予測は、天気予報や宇宙進化などの自然現象のみならず、自動車やICの設計などの工学的分野でも著しい成果を上げており、シミュレーション技術なしではこれらの分野は成り立たないことは周知の事実である。

一方、人の意思決定を総集している種々の社会現象については、一部を除いてなかなか現象を正確に再現できてはいない。それらの中で物理現象を部分的に含む道路交通などは比較的シミュレーションに成功している分野といえる。著者の属する研究グループも人が下す判断機構にファジイ推論を用いた道路交通シミュレータ(MITRAM)¹⁾を提案し、日本各地の交通状況の再現に成果を納めた。

しかし、さらに抽象的な人の行動をシミュレーションするには、相場や競りなど個々の社会現象に直接適応した専用のマルチエージェントなシミュレータを作成することが多い。しかしそのようなシミュレータを作成するまでもなく、簡易におおよその傾向を知りたいときには、汎用のマルチエージェントシミュレータが利用される。その際、使いやすさやモデル記述の自由さ、処理の速さなどといったシミュレータの要件を比較し、より自分の要望に適したシミュレータを選択・採用する必要がある。

なかでもシミュレータの使いやすさという評価軸は、モデルの理解、共有性を高めるために重要な要素である。汎用のマルチエージェント型シミュレーション環境は下記のような様々な方法で使いやすさを実現している。

(a) PlatBox Simulator²⁾：概念モデリングをサポートする機能がある

グラフィカルなモデル作成が可能(プログラミングレス)

(b) GPGSim³⁾：主に政治学・経済学分野を対象とする

簡単なJavaによる記述で様々なモデルを少ない労力で記述

(c) SOARS⁴⁾：モデルをスポットに依拠して構築(スポットオリエンテッド)

GUIでモデル構築(プログラミングレス)

(d) Repast⁵⁾：多様なモデル作成ツール(フローチャート, システムダイナミクス)

多くの雛形となるモデルを提供, 細部はGroovyやJavaで記述

(e) MASON⁶⁾：Javaによる記述でモデルを作成 ライブラリの提供でモデル作成を補助

(f) Agent sheet⁷⁾：会話型プログラミング環境 アイコンでモデルを作成

(g) NetLogo⁸⁾：多くのモデルライブラリを用意

子供の教育に使えるシンプルなスクリプト言語

(h) UFSfOM⁹⁾: 2入力1出力の要素の組み合わせでモデルを表現(著者らが提案)

GUIでモデル構築(プログラミングレス),要素から最上位モデルまでシームレス構築

このように,モデルをよりわかりやすく記述するために多くのシミュレータではGUIを用いて使いやすさを向上させる試みがなされている. UI(ユーザインタフェース)は使いやすさを評価する上で重要なもので,より使いやすさを追求するにあたってGUIの採用は必然の流れといえる.

シミュレーションのようなシステムにおけるモデルは,モデル化対象への「ある人間にとっての,ある状況,あるいは状況についての概念の明示的な解釈」¹⁰⁾である.モデル作成においては,このように解釈したモデルを如何にしてシミュレータの持つ機能を使って実現するかが問題となる.モデルの実現方法を試行錯誤する時間は実際の作業の時間と頭の中で考える時間とからなり,作業時間だけではなく思考時間の多寡にも注目することでモデルの実現容易性を評価することができることが予想される.このモデルの実現を如何に容易に行えるかという観点でシミュレータを選択する場合,シミュレータごとのモデルの実現容易性を比較評価できると便利である.

そこで本論では,シミュレータによる使いやすさの評価について実験・考察し,新しい評価方法について提案する.

2. モデルの実現容易性の評価

2.1 シミュレータの評価手法

シミュレータの使いやすさには,UIに対する評価とモデルの実現容易性がともに影響を与えており,シミュレータの使いやすさの度合いを一元的に測ることは難しい.

複数のシミュレータを比較している文献も存在し¹¹⁾,処理速度やどのような機能を持っているかという点を中心に考察されているだけである.しかし,このような比較はあくまでもシミュレータの機能評価に留まり,モデルの実現容易性の評価はなされていない.モデルの実現容易性を定量的に評価する方法を用意し,ユーザの主観的な評価と比較する必要がある.そこで本研究ではモデルの実現容易性についてユーザビリティの評価手法を利用して評価することを考えた.

ISO9241-11ではユーザビリティを「特定の利用状況において,特定のユーザによって,ある製品が,指定された目標を達成するために用いられる際の,有効さ,効率,ユーザの満足度の度合い」と定義している.この中で「有効さ」はさらに課題の完成度を評価する

完全性と、課題の成果物に含まれるエラー率を評価する確実性の二つに分けられる「効率」については作業時間や工程数で評価することが一般的であり、「満足度」については主観的評価を用いて評価することが多い。

作業時間を評価するパフォーマンス評価以外に、① アンケートを用いた主観評価をおこなう、② 脳波や心拍など生理指標の測定をおこなう、③ 実際にシミュレータを利用してもらい、その工程を評価するユーザテストを行いログを解析する、などの手法が存在する。

ユーザビリティを定量的に評価する手法に GOMS (Goal/Operator/Method/Selection rule) 法¹²⁾がある。GOMS 法に KLM (Keystroke Level Model)¹³⁾ 手法を組み合わせると Operator を分析することで、具体的にタスク遂行時の Operator 系列、実行時間、学習時間を予測し、検証・比較することができる。さらに派生として、実行時間だけではなく手法の習得時間も定量化して予測できる Natural GOMS Language や同時並列処理されるタスクに対応した CPM-GOMS などがある。

2.2 本研究での評価手法

図 1 は GOMS 法による作業時間の例を表している。

図 1

ここでは指定の場所に文字を入力するという課題に対し、課題を Goal として認識しその Goal を達成するための Sub Goal を展開している。そして Selection Rule に従ってそれぞれの Sub Goal を達成するための Method を決定している。例えば Sub Goal 1 では対象となる箇所を“クリックする事で選択する”という Method 1 を採用し、Method 1 に必要な Operator 系列を呼び出し、実行している。このとき Goal の認識、Selection Rule による Method の採用といった部分により高度な判断を含む思考時間があてられ、Method の呼び出しと Operator の実行の部分において実際の作業のための認知や動作が行われる時間があてはまる。

シミュレーションのモデル作成のような複雑な思考や判断を要求される作業に対しても Goal と Selection Rule の工程をより詳細に分析する事によって GOMS 法やその発展系を用いて解析することは可能であろう。しかしそのような解析には非常に手間がかかり、また解析者によってどのような作業がどのような思考過程で行われるかの判断が異なる事が予想されるため、GOMS 法単独による解析は適切ではないと考えられる。

そこで率直なユーザの使用感を比較的計測しやすい値で定量的に計る事を考えた。複雑な解析をせずに得られる値として、一般的に用いられる作業効率の指標である総作業時間に加え、Operator 作業にかかった時間、そして総作業時間から Operator 作業時間をひい

た残余時間の3つを用いて評価を行った。総作業時間は一般的なユーザビリティ評価の指標として知られているのに対し、Operatorの作業時間は実際にモデル作成者が行った作業量となり、残余時間はモデル作成作業時に考える時間として消費された時間となる。これらの値の中からよりユーザの主観的評価に沿った指標の探索を行った。

本研究では前述の(c)、(d)および(h)の3つのシミュレータを対象としてモデル作成をおこなうユーザテストを実施した。次にGOMS法の枠組みを参考にKLM手法を利用してOperator動作そのものにかかった課題遂行時間を推定した。

実験での課題開始から終了までの時間である総作業時間と作業内容はログ解析によって算出し、Operator作業時間をGOMS-KLM法を利用して算出した。本来GOMS-KLM法では最適なOperator系列をあらかじめ予想しその組み合わせで作業内容を計算するが、本研究では実際に観測されたOperator作業をKLM法にあてはめて評価することでOperator作業の動作にかかった時間を推測している。複雑な作業では思考と動作の時間が混在してしまい、作業ログから思考時間と動作時間を切り分けることが難しい。しかし実際の作業をもとにKLM法でOperator作業時間を推定する事で、複雑な作業であってもその作業間で行われた思考の時間を内容を包括的に取り出すことができるようになる。

ここで、実装言語の違いもまた作業時間とモデルの実現容易性に影響しているものと考えられる。プログラムコードによる記述や独自の言語による記述ではMethodそのものが異なり、その違いはKLM作業時間へと反映されている。また実装言語の違いはモデル実現のために必要な方法論にも影響し、思考時間へと反映されている。

本実験ではStep2において参考となるモデル作成を経験させることでMethodへの影響を抑えており、その上で作業内容の相違はKLM作業時間へ、モデル実現容易性の相違は思考時間へと反映されるものと考えられる。

Goalはモデル作成にあたって要求される仕様という形で事前に課題を提示した。Methodに対しては習熟とともに操作をおこなうための手続き的知識の呼び出しへの必要時間が減少していくことがわかっており¹⁴⁾、事前に必要とされるMethodについて習熟する段階を用意することで影響を最小化することを試みた。

3. 実験方法

3.1 評価実験対象のシミュレータおよび実験環境

前述のように(c)、(d)、(h)の3つのシミュレータについて評価を行った。対象となるシミュレータは被験者の負担を考慮して、実装言語による大枠での分類としてプログラムコー

ドによってモデルを完全に記述できるものとプログラミングレスかつ GUI によってモデルを記述できるものから選択した。

プログラムコードによってモデルを完全に記述できるものとして (c) の Repast Symphony2.0 の Java 開発環境を採用した。(d) の SOARS4.1.1 と (h) の UFSfOM は、共に独自規格の GUI によってプログラミングレスにモデルを作成する。(h) は著者らが提案したモデルをよりわかりやすく実現することを目標としたシミュレータで、多段二項ファジイ推論を採用することで、ユーザのヒューリスティックをそのまま取り込んだモデル作成が可能である。(a) の PlatBox Simulator もグラフィカルなモデル作成が可能で、非常に多機能かつ高性能なシミュレータである。しかし、プログラミングレスという枠組みでは SOARS 等と同様であり、初心の被験者の負担が大きくなる事をさけるため今回は対象としていない。より有意義なデータを得るために、今後これらの代表的なシミュレータに対しても実験を拡大していく必要がある。

実験はいかなるシミュレータのモデル作成の経験もない 20～30 代の技術系男性 5 名に対して行った。実験時には被験者 1 人が実験用パソコンの前に座ってもらい、験者はその隣で計測を行った。また実験中の実験用パソコンの作業を録画し作業ログの解析を行った。

被験者にはどのシミュレータがどの製作者によるものかは明かされていない。またモデル形式そのものに対する慣れによる影響を考慮して実験をおこなうシミュレータの順番はランダムとした。モデル作成時には作業時間と正確さのどちらを重視するかは被験者の任意とした。作業を中断、休息している時間は別途計測し、総作業時間から差し引いている。

3.2 実験手順

実験は 3 つの段階に分けて実施した。

Step1：サンプルモデルの内容を理解する

Step2：サンプルモデルを複製し、各シミュレータに習熟する

Step3：サンプルモデルを参考に課題モデルを作成し、ログを解析する

Repast については、Step2 では Java のコードをすべて打ち込んでもらい、Step3 では必要に応じてモジュールごとコピーすることを許可した。

実験開始前に Step1～3 のそれぞれの作業および評価手順について説明を行った。

Step1 ではサンプルとなるモデルについて自由に操作してもらい、モデルの内容を理解したと自身で判断するまでの時間を計測した。内容の理解にあたっては、モデルの入力・処理・出力構造の 3 点について理解できたかどうかを基準とした。

Step2では Method に対する習熟のためサンプルモデルの複製を行ってもらった。まったく習熟無しで作業を行った場合モデルの作成作業そのものが開始できない可能性があったためである。Step2の作業は Method への習熟度が原因となるばらつきを抑えることを目的としている。

Step3ではサンプルモデルを参考にしながらより発展したモデルを構築してもらった。Step3の課題は、5つのエージェントが持つ6つの属性をもちいて4つの対象を評価し、その中から1つを選択するもので、サンプルモデルに用いられている Method の応用のみで実現できる内容とした。被験者自身が課題モデルの作成を完全に終了したと判断した時点で作業を終了してもらった。モデル完成以前の確認・修正は可としたが、モデル完成以降の成果物に対するデバッグは行わないものとした。

Step3 終了直後、作成したモデルの動作を確認する前に 10 段階の単極尺度によって「モデル作成作業が満足にできたか」を主観評価してもらった。被験者が成果物の完成度を確認する前に評価してもらうことで作業への主観的な感想を得る事を目的とした。

4 実験結果と評価

4.1 KLM による Operator の分析

Step3 について作業内容の録画データをもとにログ解析をおこなった。ログ解析では各作業で経過した時間やマニュアルやヘルプを参照した時間などを計測することはできるが Operator の途中で随時挿入される思考の時間を取り出すことが難しい。そこでログの作業内容をもとに KLM に準拠して Operator 作業に費やされた時間を算出した。この値を KLM 作業時間と呼ぶこととする。GOMS-KLM 法によって予測する作業はエラーを含まないものであり、また想定されるユーザーは対象の作業に対して熟練しているものと仮定される。本実験では純粋に作業にかかった時間の指標とするため作業中に発生したエラーも含めたすべての実際に行われた Operator を KLM 作業時間として計算している。各処理にかかる時間は表 1 のように定義した。

表 1

KLM の各 Operator 作業時間は実験時の条件に応じて変化するが、本研究では指標としての KLM 作業時間の利用を考えたため、標準的とされる時間を採用して操作にかかった時間を推定した。

録画データの具体的な作業工程をもとにその Operator を KLM で解析、各 Operator の時間を合計することで KLM 作業時間を算出した。モデル作成の方法論は人による個性が存在するため最短の Selection Rule でない場合でも排除していない。

4.2 満足度と作業時間の比較

得られた各指標について、満足度と総作業時間(図2)、満足度とKLM作業時間(図3)、満足度と残余時間(図4)の比較を行った。ここで得た満足度は作業に対する満足度であり、製品の使用感に対する満足度とは異なる。その結果、総作業時間との相関は $-0.038(p = 0.890, R^2 = 0.000)$ で確認できなかった。KLM作業時間との相関は $0.208(p = 0.455, R^2 = 0.043)$ であった。それに対し残余時間との相関は $-0.674(p = 0.005, R^2 = 0.454)$ と負の相関が認められた。

図2

図3

図4

主観的な評価である満足度との関係を見た場合、総作業時間では評価が難しい事がわかった。さらにKLM作業時間を見た場合総作業時間よりは大きな相関が得られたが有意ではなかった。残余時間では残余時間が短いほど満足度も高くなる関係が確認され、残余時間を構成する課題達成のための思考時間がユーザの満足度を評価する指標となる可能性が示唆された。

満足度には思考時間のほかにも作業の確実性・完全性、作業時間の長期化に伴うストレスも影響していると考えられる。そのため今回はあまり強い相関を得ることはできなかったと考えられる。しかし総作業時間では主観評価とのずれが生じ評価が難しかった複雑な思考判断を主とするモデル作成課題に対して、作業工程による時間を排除した残余時間に着目することで、モデルの実現にかかわる思考時間による影響がより反映された結果を得られた。

5 各シミュレータ間での比較

5.1 主観評価による比較

主観評価によって(a)、(b)、(e)の3つのシミュレータについてモデル作成作業がどれだけ満足にできたかを評価した(図5)。(e)は著者らが提案したシミュレータであり、ユーザのモデル作成作業をどれだけ補助できているのかを他の2つのシミュレータと比較検証した。

図5

本実験ではデータ点数の確保が難しく検定は十分に信頼できる結果を得られない可能性があるためノンパラメトリック検定を採用し、Friedman検定を行ったところ全体での有意差($p < 0.01$)があった。さらにScheffeの多重比較を行い、Repast-UFSfOM間($p < 0.01$)で有意な差が見られた。モデル作成作業への満足度ではUFSfOM、SOARSが高くRepastは低くなる結果が得られた。

5.2 総作業時間とKLM作業時間および残余時間の比較

図6はStep3のUFSfOMと他の2群の総作業時間の差を示している。総作業時間は一般的

に使われる作業量の評価指標であり，作業全体の効率性を評価することができる．SOARS は最も時間がかかり，平均で 3005.2 秒，Repast は平均 1263 秒 UFSfOM は平均 1181.4 秒の時間がかかっている．

図 6

3 群に対して Friedman 検定を行ったところ，全体での有意差が認められた ($p = 0.022$)．Scheffe の多重比較を行った結果，SOARS-UFSfOM 間 ($p = 0.040$) で有意な差が見られた．この結果では 3 群の内 SOARS に比べて UFSfOM の方がモデル作成作業の効率性が高かった事になる．

さらに Step3 の KLM 作業時間では SOARS が平均で 2785.7 秒，Repast は平均 723.6 秒 UFSfOM は平均 1017.1 秒であった．Friedman 検定を行ったところ全体での有意差が認められ ($p = 0.015$)，Scheffe の多重比較を行った結果，SOARS-Repast 間において有意な差 ($p = 0.017$) が見られた．図 7 は KLM 作業時間の UFSfOM との差を示しているが，SOARS はいずれもより多くの時間が必要とされており，Repast では Subject 2 を除いて UFSfOM よりも短い時間で KLM 作業を行っている．

図 7

この結果から SOARS は同じ課題を達成するために，より多くの時間がかかっていることがわかる．しかしその時間のほとんどは KLM 作業時間に費やされたものであり，今回の課題が特別に SOARS にとって時間がかかるものであった可能性も考えられる．

そこでさらに残余時間についてみると，SOARS が平均 219.4 秒，Repast では平均 539.3 秒，UFSfOM では平均 164.2 秒であった．3 群の残余時間に対して Friedman 検定を行ったところ，有意な差は見られなかった ($p = 0.246$)．図 8 の Step3 の UFSfOM との残余時間の差を見ると有意な差はないものの UFSfOM の残余時間は他の 2 群と比較して少ない傾向にある．

図 8

これらの 3 群について定性的に見た場合，総作業時間では差があったものが，残余時間では有意といえる差がなくなっていることがわかる．総作業時間の長さは作業の効率性として無視し得ない指標であるが，SOARS では総作業時間の長さで評価した場合に反し残余時間での評価ではよい結果が得られており，モデル実現容易性は高いと考えられる．

一方で Repast では総作業時間そのものは短いにもかかわらず残余時間では多くなる傾向があった．Repast では Step3 において被験者の判断でサンプルモデルからのコードのコピーを許可したため総作業時間が短くなったものと考えられる．

UFSfOM は事前の予想では GUI の採用によって総作業時間が増大する事が想定されたが，総作業時間の平均，残余時間の平均ともに SOARS に次いで低い数字が出ている．残

余時間では他の 2 群と比べて有意な差が見られず，他の 2 群と同等のモデル実現容易性があることがわかった．

5.3 有効性に対する分析

Step3 での成果物に対して表 2 のように 6 つの達成項目を設定しそれらに対する達成度を評価した．各項目に対して確実性 2 つ，完全性 2 つの 4 点をチェックし，モデル全体が完全でなかった場合でも各項目について条件を達成していれば達成度を与えた．逆に各条件に対して成果物の中に一箇所でも未達成の箇所があればその部分の達成度を 0 とした．その後全体に対する達成数の点数を割合化して比較した．

表 2

完全性と確実性についての評価を図 9 に示す．完全性について Friedman 検定を行ったところ全体では有意差 ($p = 0.014$) があり，さらに Scheffe の多重比較を行った結果，Repast-UFSfOM 間で有意な差 ($p = 0.028$) が見られた．確実性について Friedman 検定を行ったところ全体では有意差 ($p = 0.022$) があり，Scheffe の多重比較を行った結果 SOARS-UFSfOM 間 ($p = 0.040$) で有意な差が見られ，Repast-UFSfOM 間 ($p = 0.086$) で有意傾向がみられた．UFSfOM ではモデルの不足，錯誤ともに少なく，他の 2 群ではそれぞれ 80% 程度の達成率であった．先述の残余時間では 3 群間に明確な差が見られなかったが完全性と確実性には有意な差が見られることから，同じ程度の残余時間でも作業内容の質的差があったものと考えられる．

図 9

総作業時間では SOARS が長い時間を必要とし，他の 2 群は比較的短い時間で作業を終えていた．さらに有効性の面においては，UFSfOM が最も高い水準であったことから，今回行った択一意思決定モデルの作成課題に対しては UFSfOM が適していると考えられる．

今回の実験で用いた課題は作業時間の長大化や被験者によるモデルへの不理解を抑えるための単純なモデルとして択一モデルを採用した．しかし課題が異なる場合，作業時間も異なる事が予想される．様々な課題をもちいて実験を行う事でそのシミュレータがどのようなモデル作成に適しているのかを評価することができるだろう．そこで，今後さらに多くの課題でモデル作成の容易性を比較・検証する必要がある．

6. 結論

本研究では，汎用シミュレータでのモデル作成の容易さについて種々の検討を加えた．評価の方法としては，総作業時間，ログ解析から計算した作業時間，そしてモデル作成時の課題獲得および手法選択のための思考時間に着目した．従来の GOMS 法などの分析では複雑な思考・判断を主とする作業を評価できなかったが，本研究では思考時間に着目する事

で、よりユーザの主観的評価に近い評価を行うことが可能である事がわかった。

評価には、ユーザテストの総作業時間、KLM 作業時間、そして総作業時間から KLM 作業時間を除いた残余時間の 3 つを用いた。KLM 作業時間はログ解析によって算出された作業時間であり、残余時間は思考時間に相当する時間である。実験の結果、残余時間が短いほど被験者の主観的評価での結果がよいという負の相関が確認された。このことから残余時間を比較することで複雑な思考を伴う作業にたいしても評価を行うことができることが示された。

さらに今回対象とした 3 つのシミュレータについて有効性と効率性をあわせて評価した場合、効率性の指標である総作業時間では SOARS の評価が低かったものの、残余時間ではいずれのシミュレータでも有意に差がないことがわかった。また有効性を構成する完全性、確実性を比較した結果、UFSfOM が高い有効性を示した。このことから SOARS や UFSfOM のように GUI の採用によって総時間や工程数が増えたとしても、残余時間を比較した場合モデルの実現容易性で劣るとは限らないことが示唆された。そしてこの評価方法を用いた結果、今回行った課題に対しては総作業時間も少なく残余時間が同等でかつ有効性に優れた UFSfOM が最もよいという結果が得られた。さらに異なる課題に対しての実験を行うことで、より多面的な評価を行っていきたい。

今後の研究方針としては、今回対象としなかったシミュレータに対しても実験を行い、さらに多くの種類の課題を試行した場合の傾向を調べるとともに、今回用いなかった生理的指標を含めた多様な評価を行っていく必要がある。この実験方法は被験者の負担が重くなる傾向があるため、より多くの実験結果を得るには実験内容の改善と被験者数の増員もあわせて行っていく必要がある。

参考文献

- (1) 猪飼國夫, 本多 中二, 板倉 直明, 佐藤 淳一, 佐藤 章, 中西 俊男, 高橋 道哉: ファジィ化微視的モデルによる渋滞解析を目的とした道路交通シミュレータ, シミュレーション, 16(3), 199/208 (1997)
- (2) T. Iba, Y. Takefuji : Boxed Economy Simulation Platform for Agent-Based Economic and Social Modeling. Computational Analysis of Social and Organizational Systems 2002, Pittsburgh, USA, (2002)
- (3) Saito, M., Yamaki, H., Akiyama, E., Sejima, M., and Yohida, K., : “ GPGSiM: A New Simulation Environment for International Politics and Economics, ” The 2009 Summer Computer Simulation Conference (SCSC '09), 283/290 (2009)
- (4) 田沼 英樹, 出口 弘, 清水 哲男 SOARS : 新しいエージェントベースシミュレーション言語の開発, 情報処理学会論文誌プログラミング (PRO) 46,SIG6,pp.63/63 (2005)
- (5) Collier, N., Howe, T., and North, M. Onward and upward: The transition to repast 2.0. In Proceedings of the First Annual North American Association for Computational Social and Organizational Science Conference (Pittsburgh, PA USA, June 2003), Electronic Proceedings. (2003)
- (6) Sean Luke, Claudio Cioffi-Revilla, Liviu Panait, Keith Sullivan, and Gabriel Balan. : MASON: A Multi-Agent Simulation Environment. In Simulation: Transactions of the society for Modeling and Simulation International. 82(7):517/527.(2005)
- (7) A. Repenning: Agentsheets: A Tool for Building Domain-Oriented Dynamic, Visual Environments, University of Colorado at Boulder Dept. of Computer Science, (1993)
- (8) Seth Tisue, Uri Wilensky: Design and implementation of a multiagent modeling environment, Proceedings of the Agent 2004 Conference on Social Dynamics: Interaction, Reflexivity and Emergence, Chicago, IL, (2004)
- (9) Masato Ikai, Naoaki Itakura : User-Friendly Simulator for Open Modeling by Hierarchically Management, JACIII Vol.17 No.6, 862/871 (2013)
- (10) Brian Wilson 著, 根来龍之監訳 : システム仕様の分析学, 共立出版株式会社 (1996)
- (11) Steven F. Railsback : Agent-based Simulation Platforms - Review and Development Recommendations , Society for Computer Simulation International, Vol. 82, Issue 9 , 609/623 (2006)

(12)Card Stuart, Thomas P. Moran, Allen Newell :The Psychology of Human Computer Interaction, Lawrence Erlbaum Associates (1983)

(13)Stuart K. Card, Thomas P. Moran, Allen Newell : The keystroke-level model for user performance time with interactive systems, Communications of the ACM CACM Volume 23 Issue 7, 396/410 (1980)

(14) 中村吉宏, 加藤康之, 満永豊: ユーザの思考時間に着目した習熟度指数の提案, 電子情報通信学会論文誌. A, 基礎・境界 J79-A(2), 416/423 (1996)

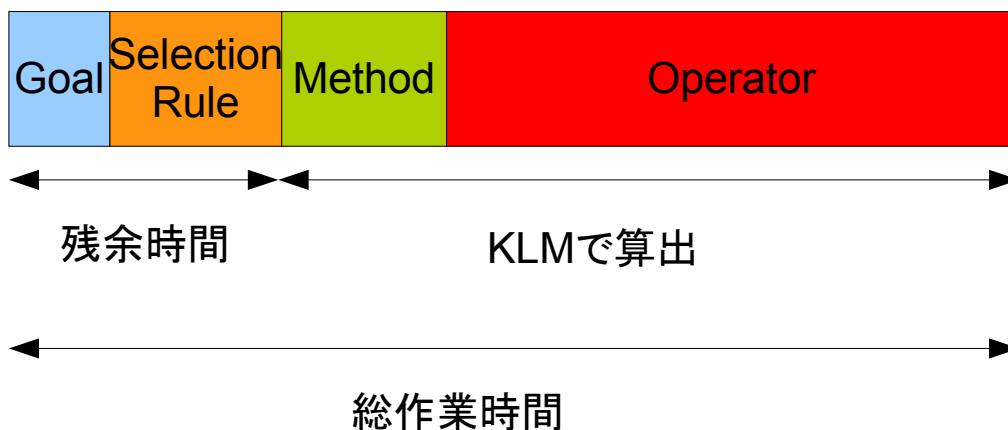
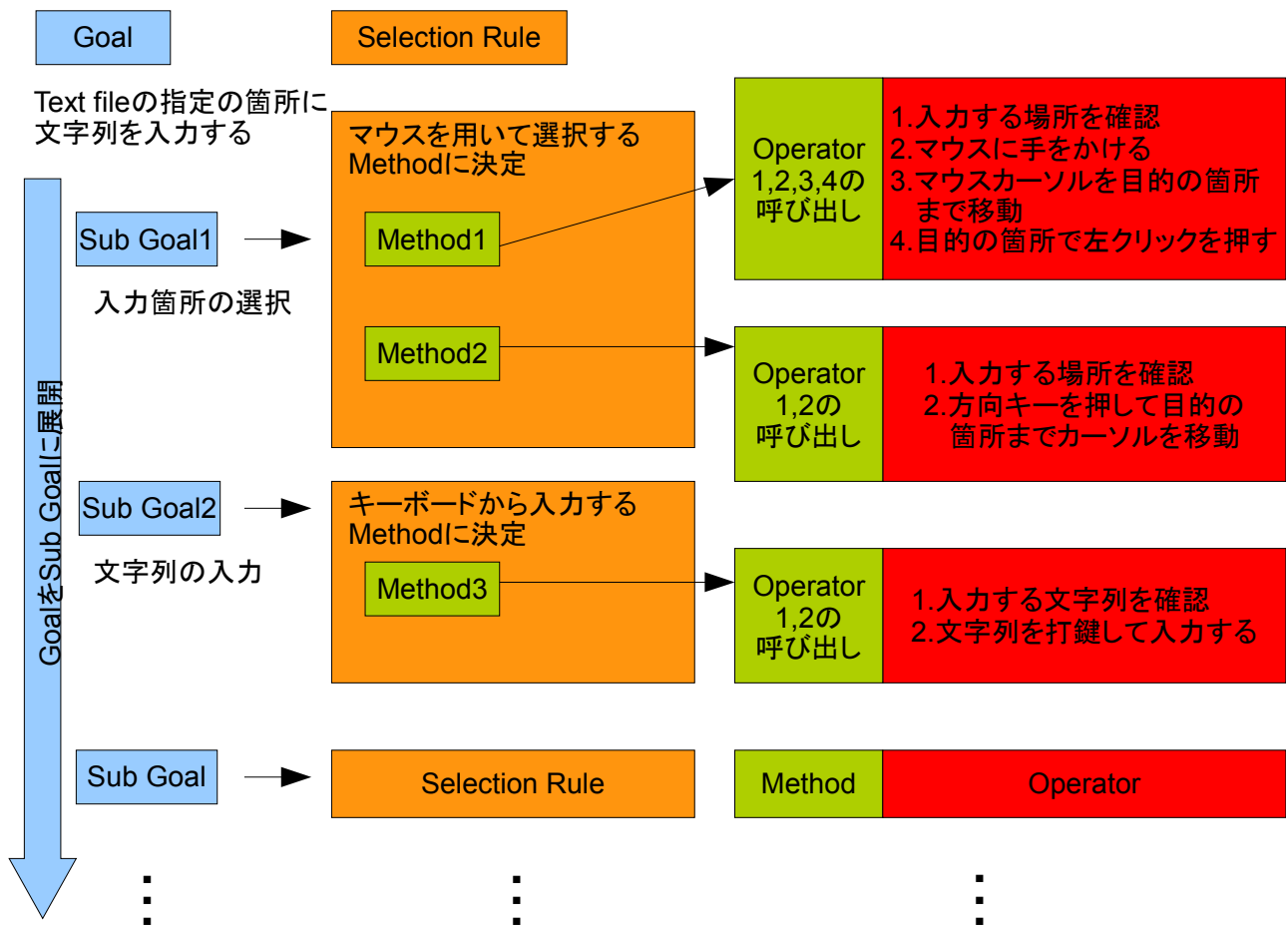


図 1 GOMS-KLM 法による作業の分析

表1 KLMにおける各 Operator の実行時間

M: 心理的準備時間(1.35sec)

H: マウスとキーボードの持ち替え(0.4sec)

P: マウスカーソルを目的の位置に移動する(1.1sec)

B: マウスのボタンを押すまたは離す(0.1sec)

K: キーボードを打鍵する(0.28sec)

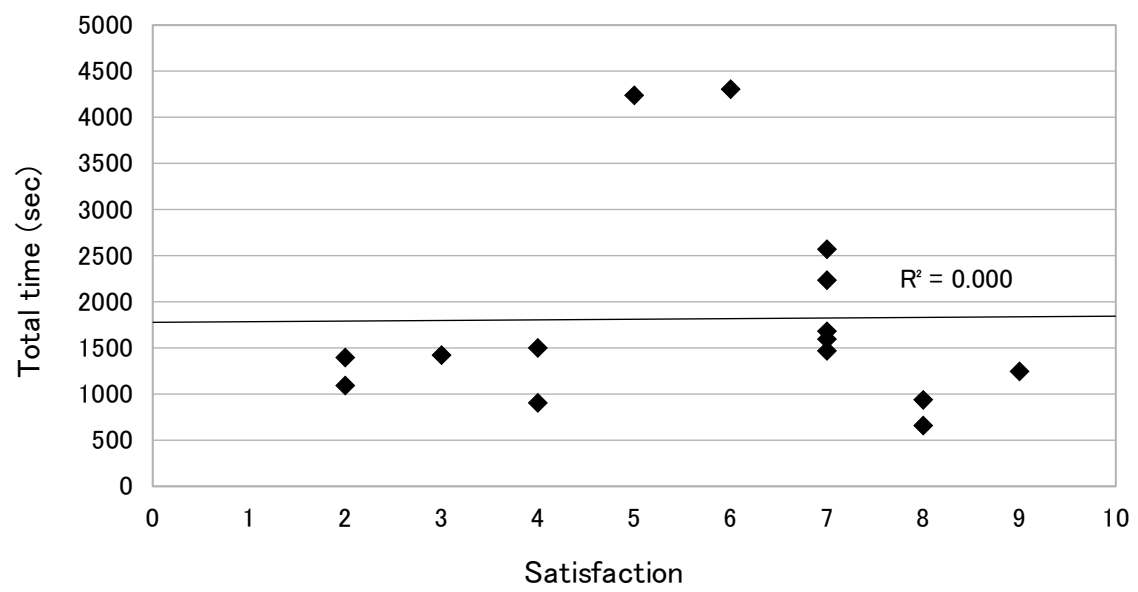


図2 満足度と総作業時間の比較

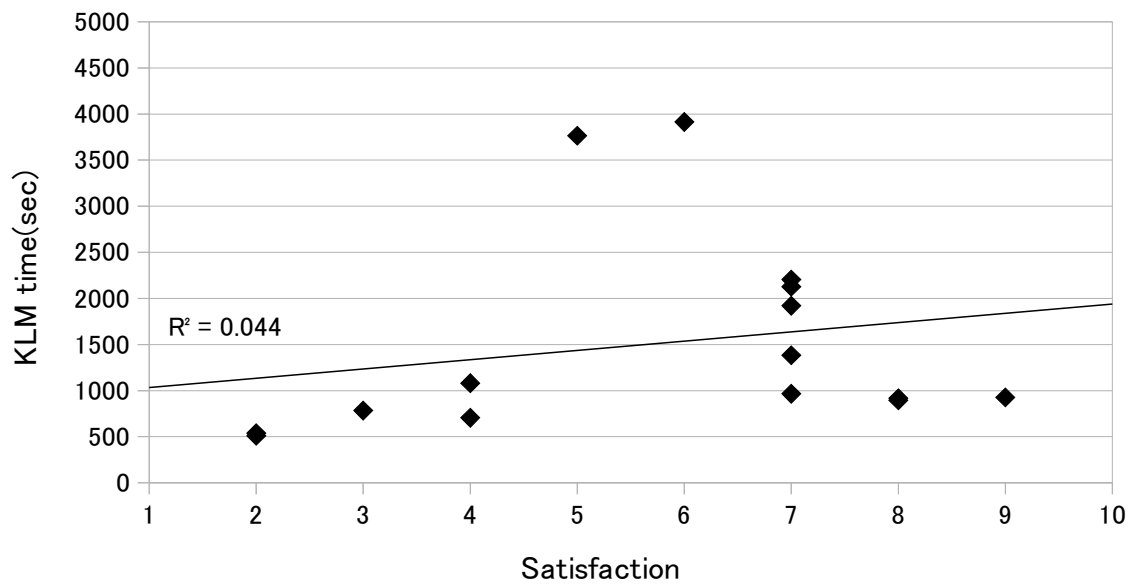


図3 満足度とKLM作業時間の比較

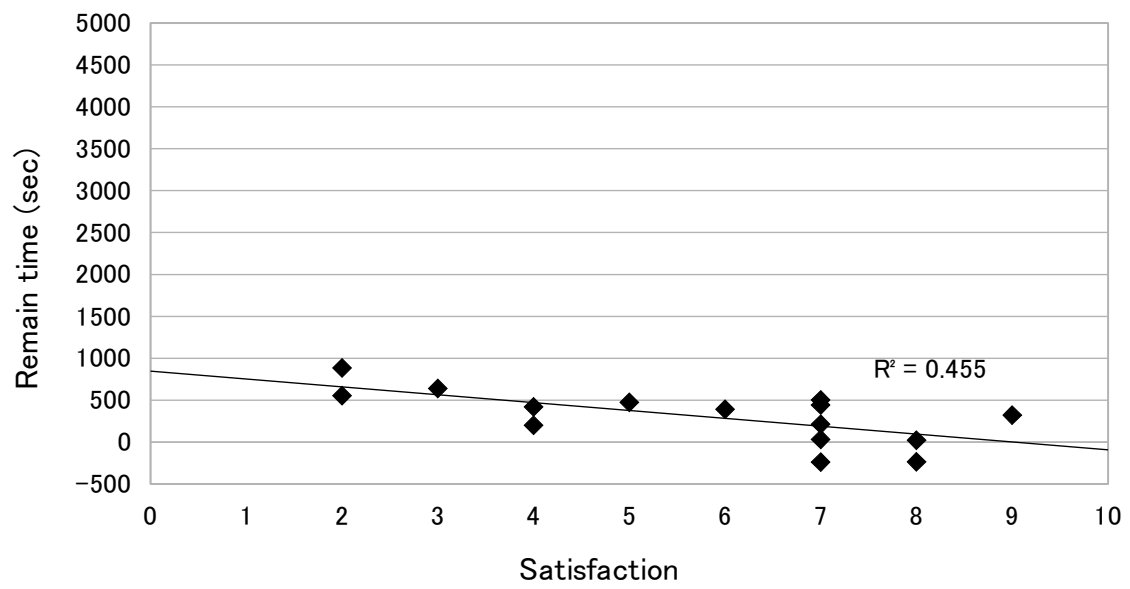
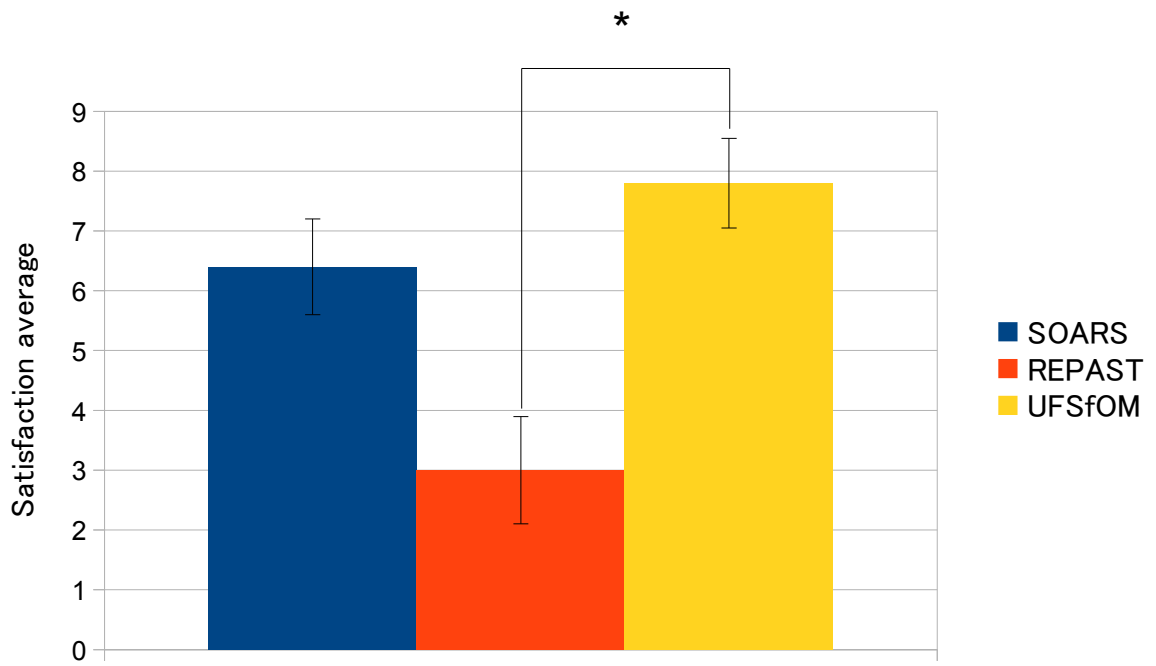


図4 満足度と残余時間の比較



* : p<0.05

図5 モデル作成作業に対する満足度

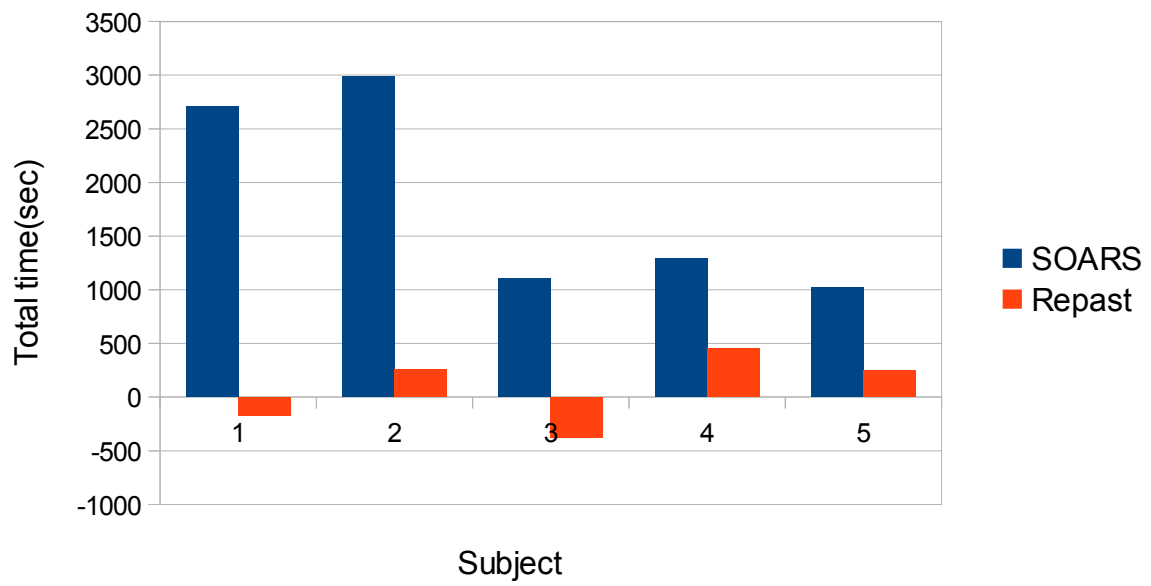


図 6 Step3 における UFSfOM と他群の総作業時間の差

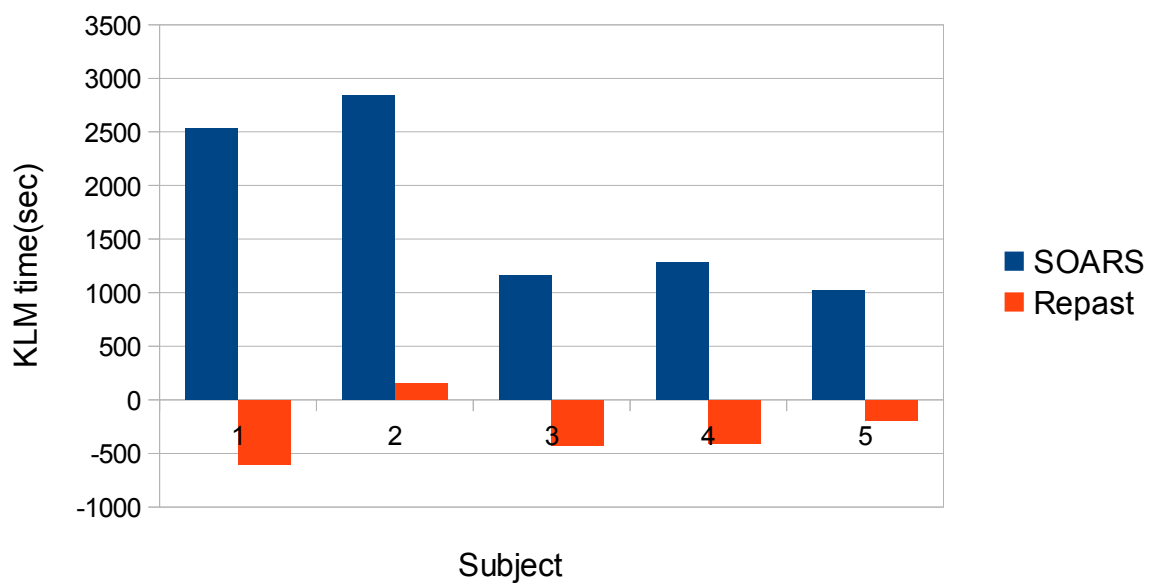


図 7 Step3 における UFSfOM と他群の KLM 作業時間の差

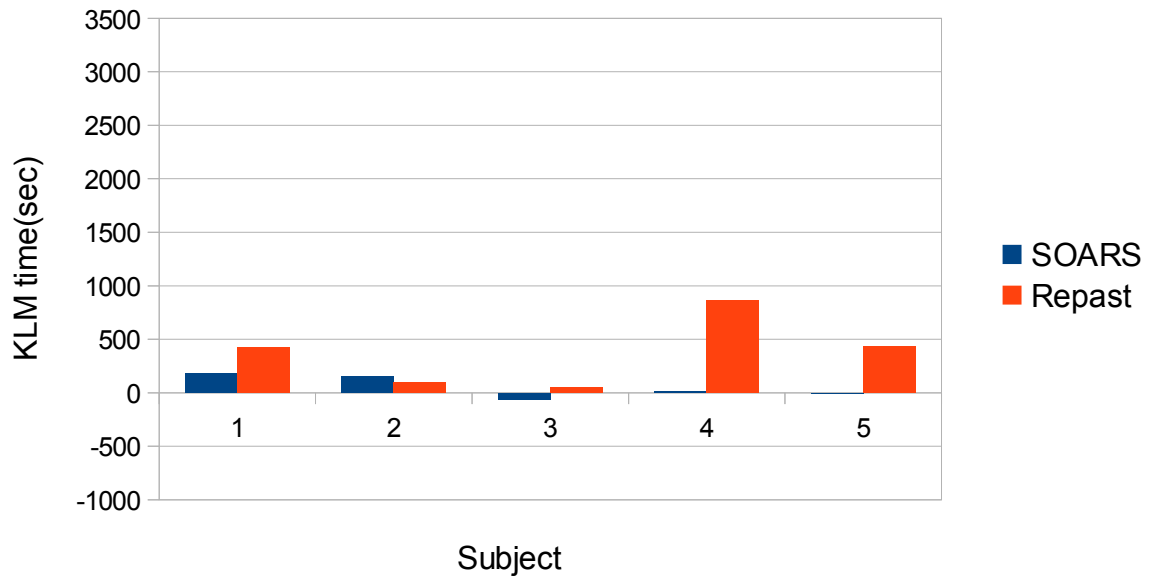


図 8 Step3 における UFSfOM と他群の残余時間の差

表2 確実性と完全性の評価項目

- ・エージェントは5つ作られているか.
- ・エージェントは値の比較を行って対象を選択できるか.
- ・値を算出するための属性は6つ設定されているか.
- ・入力される数値は正常に処理されているか.
- ・選択対象は4つ設定されているか.
- ・選択結果の出力は正常に処理されているか.

確実性:

- 選択されたMethodはモデルを実現できるものか.
- モデルの記述に錯誤はないか.

完全性:

- モデル実装に必要な記述に不足はないか.
- モデル実装に不要な記述が動作を妨げていないか.

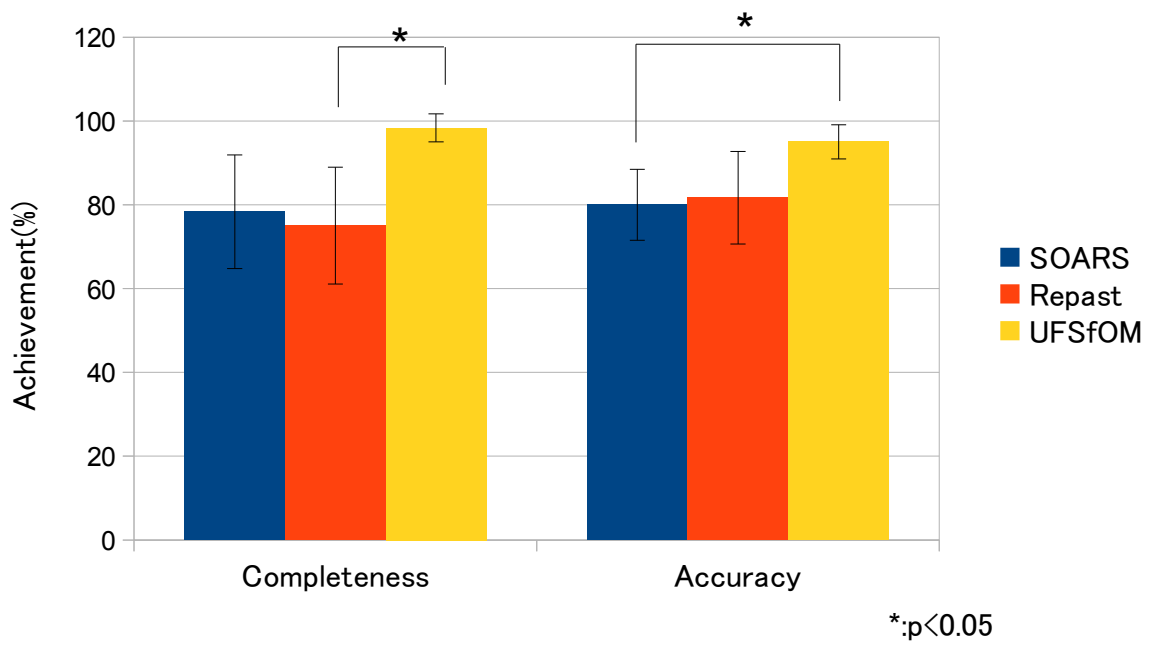


図9 3群における確実性と完全性の評価

図表一覧

- 図 1 GOMS-KLM 法による作業時間の分析
- 表 1 KLM における各 Operator の実行時間
- 図 2 満足度と総作業時間の比較
- 図 3 満足度と KLM 作業時間の比較
- 図 4 満足度と残余時間の比較
- 図 5 モデル作成作業に対する満足度
- 図 6 Step3 における UFSfOM と他群の総作業時間の差
- 図 7 Step3 における UFSfOM と他群の KLM 作業時間の差
- 図 8 Step3 における UFSfOM と他群の残余時間の差
- 表 2 確実性と完全性の評価項目
- 図 9 3 群における確実性と完全性の評価