

SOARS Toolkitを用いた社会シミュレーションプラットフォームの構築

○松井陽太郎 嶋直紀 渡邊康平 市川学 (芝浦工業大学)

Construction of a social simulation platform using the Soars Toolkit

*Y. Matsui, N. Shima, K. Watanabe, M. Ichikawa (S.I.T.)

概要— 社会シミュレーションの手法の一つとしてエージェントベースモデリング (ABM) が広く用いられている。ABMによる社会現象のモデル化はシンプルなモデル構築が主であったが、計算機資源の潤沢化によりモデルの複雑化が許容されてきている。そこで日常生活の中で行われる社会現象のシミュレーション構築を容易にするために、人口構成や都市構造を実装、出力結果の分類方法提案する社会シミュレーションプラットフォームの構築を行った。本項ではその構築方法について中心に述べる。

キーワード: エージェントベースモデル, 標準行動モデル, 合成人口

1 はじめに

1.1 エージェントベースモデル

エージェントベースモデリング (Agent Based Modeling, ABM) は社会科学で用いられる強力なモデリング手法のひとつである。ミクロな「エージェント」と呼ぶ内部状態と意思決定, 問題解決能力, 並びに通信機能を備えた活動主体を定義し, それらの相互作用を記述することにより, マクロな社会現象をボトムアップにモデル化することができる。また, エージェントは自身を取り巻く環境であるマクロな社会現象から影響を受け, 相互的な影響の及ぼし合い (ミクロ・マクロリンク, ミクロ・マクロループ) を形成することが特徴である。

1.2 ABM の設計の複雑化

社会現象のモデル化に対する議論として, KISS (Keep It Simple, Stupid!) の原則に従うことが通例とされていた¹⁾。しかし, 近年では計算機リソースの潤沢化により, より複雑で大規模なシミュレーションを行うことが可能であり, また従来の抽象化されたモデルに対して, 疑問が持たれている。

1.3 ABM 記述言語

複雑化する ABM の設計, 実装の負荷を軽減する一つのアプローチとして, さまざまな ABM 記述言語の提案が行われてきた。NetLogo²⁾, MASON (Multi-Agent Simulator Of Neighborgoods)³⁾, Artisoc (Artificial Societies)⁴⁾ はセル型と呼ばれる格子状の空間概念を採用し, エージェントの活動をその空間の移動と, 空間内での相互作用を表現することで対象となる社会現象をモデリングする。また, 出口らはエージェントが相互作用する場であるスポットという空間概念を採用した SOARS (Spot Oriented Agent Role Simulator)⁵⁾ を提案している。

空間概念の取扱方によりモデルのシミュレーション結果に影響を与えることが市川 2013⁶⁾ で指摘されている。市川 2013 では従来型の格子状の空間概念と, スポットの空間概念を用いた仮想都市モデルを構築し, 感染症シミュレーションを行った際の影響を検証した。エージェントには職業や, 年齢階級などの属性を持たせ, それぞれの属性に合わせた行動を行うよう設計している。

その結果, 格子状空間では感染症の蔓延の広がりが早いことや本来生活空間が重なり合うことのないエージェント同士が感染を広げるなどの点を指摘し, 表現する社会現象の特徴に応じて, 仮想都市モデルの構築法を考慮する必要があることを示した。

1.4 社会シミュレーションプラットフォーム

もう一つのアプローチとして, 日常的な人々の生活空間や活動を定義した社会シミュレーションプラットフォームを構築し, その上で対象となる社会現象を実装する方法を市川 2017⁷⁾ では提唱している。社会シミュレーションでは, 現実社会の人々が日常の行動を行っている中で発生する社会現象に注目した研究は少なくない。そのため市川らは職業, 年齢などを加味した標準生活行動モデルを提案し, その構築方法を示した。

そこで, 本項ではスポットの空間概念を採用し, エージェントの生活空間や標準行動を実装した社会シミュレーションプラットフォームの構築方法について述べる。また, ケーススタディとして本校のプラットフォームを用いて構築した感染症シミュレーションの結果を用い, その出力結果の分類手法について述べる。

2 方法

本研究では, エージェントベースモデルにより国民の行動, 人口および都市構造をミクロレベルでモデル化するとともに, シミュレーションの構築において, Java 言語で実装された社会シミュレーション・ライブラリ SOARS Toolkit⁸⁾ を用いて記述した。SOARS Toolkit は, SOARS (Spot Oriented Agent Role Simulator) の設計思想を引き継いでおり, スポット・ロール・ステージモデルの実装を容易にする。SOARS の設計は感染シミュレーションへの実用性が認められており, SOARS Toolkit はエージェント数 3000 万を超える首都圏のシミュレーションを想定して実装されている。また人口のミクロレベルでのモデル化には, 原田 (2018)⁹⁾ らが作成した大規模世帯合成データ, 模擬個票を用いた。模擬個票は国内の対象地域における世帯員の年齢や職業などの世帯構成を含む母集団の模擬データである。国内の複数の統計データを用い, 組み合わせ問題を解くことで全体の母集団を合成している。また都市構造の再現には, e-Stat¹¹⁾ や国土数値情報¹²⁾ から取得できるオープンデータなどを利用した。

3 モデルの構成

本研究の社会シミュレーションプラットフォームは2つの基礎モデルと、ユーザーが実装する社会現象モデル、出力結果の解析手法からなる。それぞれのモデルの説明、社会現象を実装する際の設計概念、出力結果の分類方法について説明する。

都市と人口構成モデル 住民の人口構成、及び住民のエージェントが生活する都市空間のモデル化。

多様な住民行動モデル 住民の生活をエージェントの行動としてモデル化

社会現象モデルの実装 ユーザーが実装する社会現象のモデル

出力結果の分類 大規模 log を自動的に分類する方法

3.1 都市構造と人口構成モデル

3.1.1 都市構造モデル

都市構造のモデル化では、事業所、教育機関をエージェントが相互作用する空間であるスポットとして表現した。e-Stat¹¹⁾の経済センサス-基礎調査からは、1kmメッシュ単位の全産業事業所数及び全産業従業者数が得られる。これより1kmメッシュの中に産業ごとの事業所をスポットとして複数生成した。国土数値情報¹²⁾からは全国の学校のポイントデータが得られる。すべての学校を、それぞれを含む1kmメッシュに対応づけてスポットとして生成した。生成したスポットをTable 1にまとめる。

3.1.2 合成人口データを用いた人口構成モデル

人口構成のモデル化を説明する。人口構成のモデル化には模擬個票を用いている。模擬個票には、個人データ、世帯データおよび建物データなどが含まれる。個人データに含まれる仮想個人には、個人ID、世帯ID、年齢、性別、世帯内役割などの属性がある。加えて就業先の産業分類も付加されている。また世帯データに含まれる仮想世帯には、家族類型、人数、建物IDなどの属性がある。建物IDは建物データに含まれる建物のIDであり、すべての建物には国内の住所の緯度経度が割り当てられている。そのため国内の夜間人口が再現されている。また提供いただいた模擬個票には、モバイル空間統計を用いた方法によって、通学先・通勤先の地域が500mメッシュ単位が割り当てられている。本研究のモデルの都市構造は1kmメッシュの精度で生成している。そこで1kmメッシュの範囲内に存在する事業所や学校のスポットを、それぞれのエージェントの通学先あるいは通勤先として割り当てることで、国内の昼間の人口動態を再現した。エージェントの持つ属性はTable 2にまとめる。

3.2 多様な住民行動モデル

対象地域の住民のエージェントの行動を再現するにあたって、市川ら(2017)が提案した日常の標準生活行動のモデル化を行った⁷⁾。モデル化にあたって市川らと同様にデータブック国民生活時間調査¹³⁾(国民生活時間調査)を用いた。国民生活時間調査は日本人の睡

眠や労働、家事、食事など29種の行動を、職業、年齢、性別などの分類(Table. 3)ごとに集計し、15分毎の行為率を調査したものである。

標準生活行動のモデル化は、与えられた統計情報に合うように、行為を行う人々のスケジュールを当てはめていく組み合わせ問題としてSA法で解く。

SA法のアルゴリズムの詳細を説明する。市川らの方法では10万人の行為者を用意していたが、計算時間を短縮するため、また同一スケジュールが大量生成されるため結果が冗長となることから、100人分の行為者を用意した。それぞれの行為開始時刻を $x_{i(i=1,2,\dots,100)}$ とする。初期値として1日の15分刻みの時刻を一様乱数で与える。 x_i が取り得る値は、15分刻みの1日を0から95までの整数値で表し、(0,1,...,95)の中から1つの値をとる。例えば0:00は0、0:15は1と表され、95は23:45を表す。また行為者は、ある時刻の行為を判定する96次元の配列 X_{i96} を持つ。整数で表現する時刻 j で行為を実行している場合、 $X_{ij} = 1$ 、実行していない場合は $X_j = 0$ となる。国民生活時間調査の結果(行為者平均時間量)を15分で除した整数値を用い、行為時間の初期値 s_i として与えた。そのため、 $Xx_i = Xx_i + 1 = \dots = Xx_i + s_i = 1$ となる。ただし $x_i + s_i$ が配列を超える場合、96で除した余りの整数値 r_i を求め、 $X_{ir_i} = 1$ となる。また行為者全体のみならず、ある時刻に行為を実行している割合を保持する配列 Y_{96} を定義し、

$$Y_j = \sum_{i=1}^{100} \frac{X_{ij}}{100} \quad (j = 0, 1, \dots, 95)$$

とする。評価値については、国民生活時間調査の結果(時刻別行為者率の統計値) R_j を利用し、 $E = \sum_{j=0}^{95} (Y_j - R_j)^2$ とした。加えて、学業や仕事の時刻別行為者率の統計を観察すると、昼食などの中休憩のために正午周辺で行為者率が急減し、午後からふたたび急増する様子がみられる。よって中休憩時刻 p_i を行為時間帯の中間に与え、中休憩時間 $t_i = 4$ は1時間と仮定した。そのため、 $Xx_i = Xx_i + 1 = \dots = Xx_i + p_i - 1 = Xx_i + p_i + 4 = \dots = Xx_i + s_i = 1$ となる。アニーリングについては、あるひとりの行為者の行為開始時刻 x_i を前後1時間の間でランダムに変更することで次の状態を作成する。また休憩時刻も前後15分の間でランダムに変更できるものとした。初期温度 $T = 0.01$ 、冷却率 $r = 0.99$ 、終了条件 $T \leq 0.001$ とした。また評価値の変化量 $\Delta E \geq 0$ の場合、すなわち誤差が拡大した場合の遷移確率は、

$$\exp\left(-\frac{\Delta E}{T}\right)$$

とした。Fig.1は結果の一例である。青い線が、用意した100人の行為者のスケジュールの集計値であり、灰色の線が国民生活時間調査の時刻別行為者率である。

データブック国民生活時間調査では、生活時間に特徴の表れる人々の分類毎に集計を行っており、それらの分類と模擬個票における個人の属性を対応させることで、個人にあった生活時間が再現できる。作成した対応表はTable4である。結果として、ある行為者分類に対して仕事または学業の100通りのスケジュール

Table 1: スポットの種類一覧表

スポットの種類	データ元
家庭	模擬個票における仮想世帯
小学校	国土数値情報 学校データ
中学校	国土数値情報 学校データ
高校	国土数値情報 学校データ
大学	国土数値情報 学校データ
事業所C 鉱業, 採石業, 砂利採取業	e-Stat 経済センサス-基礎調査
事業所D 建設業	e-Stat 経済センサス-基礎調査
事業所E 製造業	e-Stat 経済センサス-基礎調査
事業所F 電気・ガス・熱供給・水道業	e-Stat 経済センサス-基礎調査
事業所G 情報通信業	e-Stat 経済センサス-基礎調査
事業所H 運輸業, 郵便業	e-Stat 経済センサス-基礎調査
事業所I 卸売業, 小売業	e-Stat 経済センサス-基礎調査
事業所J 金融業, 保険業	e-Stat 経済センサス-基礎調査
事業所K 不動産業, 物品賃貸業	e-Stat 経済センサス-基礎調査
事業所L 学術研究, 専門・技術サービス業	e-Stat 経済センサス-基礎調査
事業所M 宿泊業, 飲食サービス業	e-Stat 経済センサス-基礎調査
事業所N 生活関連サービス業, 娯楽業	e-Stat 経済センサス-基礎調査
事業所O 教育, 学習支援業	e-Stat 経済センサス-基礎調査
事業所P 医療, 福祉	e-Stat 経済センサス-基礎調査
事業所Q 複合サービス事業	e-Stat 経済センサス-基礎調査
事業所R サービス業 (他に分類されないもの)	e-Stat 経済センサス-基礎調査

Table 2: エージェントの属性表

属性	説明	個人の例
name	エージェントの名前	p4
person_id	模擬個票上の個人 ID	113247898
household_id	模擬個票上の世帯 ID	49405883
age	年齢	21
sex_id	0=男 1=女	1
role_household_id	世帯における役割の ID	1(=単独世帯 (女性))
employment_type_id	就労形態の ID	20(=短時間労働者)
industry_code	産業コード	M(=宿泊業, 飲食サービス業)
student_type_id	学生タイプ ID	univ(=大学生)
city_code	市区町村コード (通勤・通学先)	46535
mesh_code	1km 地域メッシュコード (通勤・通学先)	40284344
life_style_id	国民生活時間調査における分類 ID	sales(=販売職・サービス職)

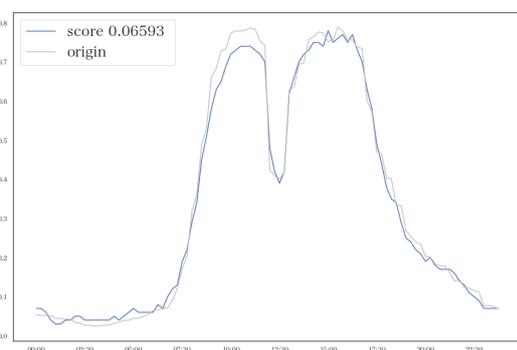


Fig. 1: 事務職・技術職平日の仕事行為者率

ジュールの一例である。縦軸が1日の時刻であり、横軸が100人の行為者である。オレンジの短形で塗られた範囲がそれぞれの行為時間帯を表している。

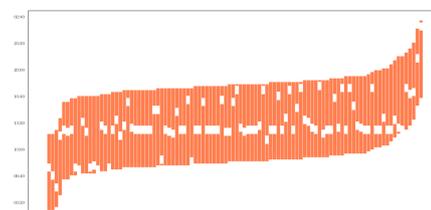


Fig. 2: 事務職・技術職平日の仕事スケジュール

が生成できた。生成したスケジュールは対象とする母集団のおよそ100分の1の人数が回答したスケジュールであると解釈できる。そのため生成した100通りのスケジュールをランダムに模擬個票の個人に割り当て、エージェントの日常的な生活を再現した。Fig.2はスケ

本プラットフォームでは、標準行動の中から1日の多くを占める通勤通学を生活者ロールとして定義し、実装した。生活者ロールは Table 5 に示す。

Table 3: NHK 分類 (一部抜粋)

日本語シート名	encode
農林漁業者平日	industry_01_weekday
販売職・サービス職平日	industry_03_weekday
技能職・作業職平日	industry_04_weekday
事務職・技術職平日	industry_05_weekday
経営者・管理職平日	industry_06_weekday
専門職・自由業・その他平日	industry_07_weekday
学生平日	industry_10_weekday
小学生平日	school_elementary_weekday
中学生平日	school_junior_weekday
高校生平日	school_high_weekday

Table 4: 個人の属性と集計対象分類の対応表

個人の属性	国民生活時間調査	行為
A 農業, 林業	農林漁業者	仕事
B 漁業	農林漁業者	仕事
C 鉱業, 採石業, 砂利採取業	技能職・作業職	仕事
D 建設業	技能職・作業職	仕事
E 製造業	技能職・作業職	仕事
F 電気・ガス・熱供給・水道業	技能職・作業職	仕事
G 情報通信業	事務職・技術職	仕事
H 運輸業, 郵便業	技能職・作業職	仕事
I 卸売業, 小売業	販売職・サービス職	仕事
J 金融業, 保険業	事務職・技術職	仕事
K 不動産業, 物品賃貸業	事務職・技術職	仕事
L 学術研究, 専門・技術サービス業	専門職・自由業・その他	仕事
M 宿泊業, 飲食サービス業	販売職・サービス職	仕事
N 生活関連サービス業, 娯楽業	販売職・サービス職	仕事
O 教育, 学習支援業	事務職・技術職	仕事
P 医療, 福祉	事務職・技術職	仕事
Q 複合サービス事業	事務職・技術職	仕事
R サービス業 (他に分類されないもの)	事務職・技術職	仕事
S 公務 (他に分類されるものを除く)	事務職・技術職	仕事
T 分類不能の産業	専門職・自由業・その他	仕事
小学生	小学生	学業
中学生	中学生	学業
高校生	高校生	学業
大学生	男20代, 女20代	学業

3.3 社会現象モデルの実装

本稿のプラットフォームへの社会モデルの実装方法とその利点について、感染症シミュレーションを例に述べる。まず、感染症シミュレーションについて定義する。新型コロナウイルスや天然痘などの感染症(伝染病)は、病気を起こした人から、病原体が別の人へと到達し連鎖的に感染現象が発生する。人と人の活動の中で感染現象がどのように拡大していくのか、また対策の効果を検証するためのシミュレーションを感染症シミュレーションとする。本プラットフォームでは、国勢調査を元にした、小地域で集計した場合正しい人口と、国民生活時間調査を元に再現した年齢、性別、職種に基づいた生活者ロールを再現している。感染症シミュレーションの実装には、感染現象のプロセスと感染症の病態遷移を組み込むことで、感染症のシミュレーションをモデル化し、組み込むだけで良い。

3.4 出力結果の分類

3.4.1 大規模エージェントのログの解析

ABMを用いたシミュレーションは、モデルの有用性の検証や対策シナリオの効果を検証するため、シミュレーションを複数回も実行する必要がある。エージェント一人一人を追跡可能という利点を持つ反面、出力されるログの量が大量になってしまうという欠点がある。そこで出力されるログの傾向を自動的に分類する機能を実装した。

3.4.2 エージェントのステータスに着目した分類手法

シミュレーションを実行した際のエージェントのステータスの出力結果は Table 6 のようになる。都市構造と人口構成モデルの生成元となる模擬個票に同一のものを使用すれば、複数回のシミュレーションに渡り、エージェントは同一のエージェントであることが担保される。そのため、エージェントの位置、隣接関係と

Table 5: 生活者ロール

ルール名	説明	ステージ
移動ルール（通勤・通学）	家庭から通勤・通学先のスポットへ移動	移動ステージ
移動ルール（帰宅）	通勤・通学先から家庭のスポットへ移動	移動ステージ

ステータスの変化はシミュレーションの結果を表す特徴量と見なすことができる。この特徴量行列の類似性を用いてエージェントのログの分類を行う。

3.4.3 k-means 法を用いた分類結果

先述の特徴量行列に対して機械学習の教師なし学習で行われる分類手法の k-means 法を適応する。k-means 法は行列同士の類似度を示す距離関数を用いて、指定された分類数でクラス分けを行う分類手法である。分類器の実装は python の機械学習ライブラリ scikit-learn ライブラリ (sk-learn) を用いて行っている。sk-learn で実装された k-means 法はユークリッド距離を用いて行列の計算する。ログの解析にあたり、シミュレーション結果を時間軸を全て連結した 1 次元の配列に変形している。

Table 7 は本稿のプラットフォームを用いて構築した感染症シミュレータを例にして、9 回分のシミュレーション結果に対してクラス分類を行った結果である。

k-means 法の分類結果は別のカテゴリであるということを示しているだけであり、同一の数値、連続性に対しては意味を持たない。出力結果から、分類には一定の法則があり、分類の数によってある程度の意味解釈を持つことができる範囲を規定することができそうである。上記の分類結果から class 5, class 6, class 7 で分類したの感染者の推移状況を示す。Fig.3:,Fig.4:,Fig.5:

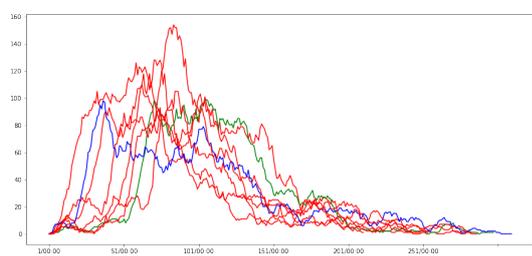


Fig. 3: :5 クラスの分類結果

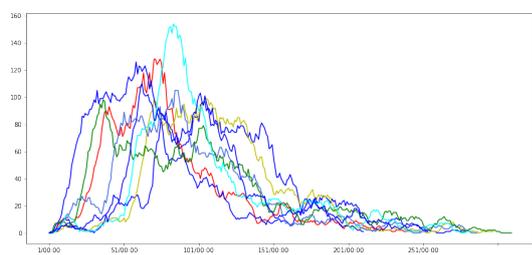


Fig. 4: :6 クラスの分類結果

同一の色が同一をクラスであることを示している。class 5 と class 6 を比較した場合、感染者数が大きく飛び出している部分の log を分類できている。

4 まとめ

本項では ABM の設計、実装、シミュレーション結果の解析の負担軽減のための社会シミュレーションプラッ

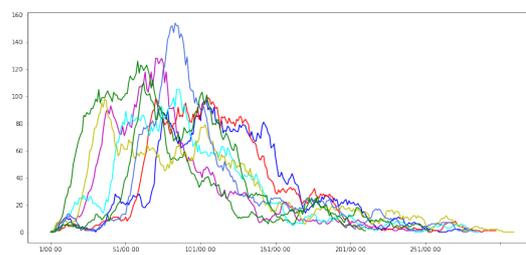


Fig. 5: :7 クラスの分類結果

トフォームの提案を行った。本プラットフォームを用いることで、日常生活と密接な関係を持つさまざまな社会現象のモデル化が容易になると期待できる。出力ログの解析結果は、現在あまり精度の高いものでないこと、またそれぞれの対策シナリオに一樣の結果を示せるかなどの課題が残っている。大規模なシミュレーションの結果を容易に理解する手法を同時に確立することにより、分類結果の優位性の確認を行うことができるのではと思うため、今後の課題とする。

参考文献

- 1) 寺野 隆雄：なぜ社会システム分析にエージェント・ベース・モデリングが必要か、横幹, 4 号, 56/62 (2010), 公開日 2016/03/21, Online ISSN 2189-6399, Print ISSN 1881-7610
- 2) <https://ccl.northwestern.edu/netlogo/>
- 3) <https://cs.gmu.edu/~eclab/projects/mason/>
- 4) <https://mas.kke.co.jp/artisoc4/>
- 5) Tanuma H., Deguchi H., Shimizu T. (2005) SOARS: Spot Oriented Agent Role Simulator — Design and Implementation. In: Terano T., Kita H., Kaneda T., Arai K., Deguchi H. (eds) Agent-Based Simulation: From Modeling Methodologies to Real-World Applications. Agent-Based Social Systems, vol 1. Springer, Tokyo.
- 6) 市川 学, 出口 弘：感染症実用シミュレーションにおける仮想都市構築法の違いによる結果への影響分析. 計測自動制御学会論文集. **49** 巻 11 号, 1012/1019 (2013) .
- 7) 市川 学, 小森賢一郎：人々の日常の標準生活行動モデルの提案に向けた統計情報分析, SICE 社会システム部会, <http://journals.socsys.org/symposium012/pdf/012-024.pdf>, 参照日 2020/07/01.
- 8) 小野, 市川, 出口：大規模エージェントベースシミュレーションのための SOARS Toolkit の提案, SSI2020, GS6-4-5 (2020)
- 9) 原田 拓弥, 村田 忠彦, 柘井 大貴：家族類型と世帯内の役割を考慮した SA 法による大規模世帯の合成, 計測自動制御学会論文集, **54** 巻 9 号, 705/717 (2018) .
- 10) <https://mobaku.jp/>
- 11) <https://www.e-stat.go.jp/>
- 12) <https://nlftp.mlit.go.jp/ksj/>
- 13) NHK 放送文化研究所：データブック国民生活時間調査 2015

Table 6: エージェントの時刻におけるステータス

time	agent1	agent2	agent3	...
0/00:00	0	1	0	
0/01:00	1	2	0	
...				

Table 7: k-means 法によるクラス分類

seed	class3	class4	class 5	class 6	class7	class8
agentstate-6379524053489412280	0	1	0	2	0	4
agentstate-5804828401988331612	0	1	0	0	1	7
agentstate629573471579609896	1	0	4	3	5	3
agentstate-5637448352339666157	0	1	0	2	2	0
agentstate9119843386004838872	0	1	0	2	2	6
agentstate7350630895459360843	0	2	3	5	3	1
agentstate-7609672752176747469	2	3	2	1	4	2
agentstate7943734622742647041	0	1	1	4	6	5
agentstate-4083129752674054010	0	1	0	2	2	0