

## 行動素のワークフローへの自動対応付け手法

楼天甲<sup>†</sup> 江川陽<sup>†</sup> 白山晋<sup>††</sup> 稗方和夫<sup>†</sup> 大和裕幸<sup>†††</sup>

行動分析にもとづくワークフローの詳細化において、行動とワークフローの対応付けが問題になる。本稿では、計測によって得られた行動素のワークフローへの自動対応付け手法を提案する。

### Automatic Alignment of Motion Features to the Workflow

TIANJIA LOU<sup>†</sup> AKIRA EGAWA<sup>†</sup> SUSUMU SHIRAYAMA<sup>††</sup>  
KAZUO HIEKATA<sup>†</sup> HIROYUKI YAMATO<sup>†††</sup>

Workflow has become one of the key components to representing engineering knowledge. And, detailed specifications of the workflow have been shown to be effective for extracting knowledge from experts. In our previous work, we proposed an approach for detailed specifications of the workflow based on human behavior analysis, and showed that the approach becomes one of the techniques for effective knowledge extraction. However, it is found that extracting motion features from sensing data and aligning those to the workflow become the bottleneck in our approach. In order to resolve the bottleneck, we propose a new approach of automatic alignment of motion features to the workflow.

#### 1. はじめに

操業現場において、システムと操業者のやり取りや、操業者自身の行動に着目すると、操業に関連する有用なデータが自然な形で日々大量に生じていることがわかる。これは利用者からのフィードバック情報や、熟練者のピアレビューに相当するものと考えられる。

このような人間の行動によって生じるデータから形式知を抽出するという試みがある。操業者の知識として記述されていない慣習的な行動や、気づきを顕在化するという試みでもある。

Harper は、熟練工の作業の様子に着目し、その様子を撮影し、インタビューによって分析を試みた<sup>1)</sup>。Kimuraらは、熟練設計者のCAD操作画面を動画によってキャプチャし、手順や判断基準をワークフローによって形式知化している<sup>2)</sup>。成子らは、Kimuraらの知見にもとづき、ワークフローをマクロフローとミクロフローに分類し、以下の手順で知識のデータベース化を試みている<sup>3)</sup>。

- 熟練者の有する知識やノウハウを設計・製作の作業手順毎に分類する。
- 第三者に伝達可能な図表や文章で表現する。
- (b)を関連情報とともに、フローチャートに数表、図表、具体的指示などを関連付ける。
- (c)をデジタル情報としてデータベース化する。

Hiekataらは、ヒアリングによって初期ワークフローを生成し、そのワークフローに相当する実作業を分析することでワークフローの修正を行い、初期ワークフローと修正後のワークフローの差から知識抽出を試みている<sup>4)</sup>。

これらの研究はワークフローを介した形式知化であるが、行動自体の分析によって慣習的な行動や、気づきを顕在化するという試みがある。行動に対するセンシングデータの取得とその分析が一例である。

石油精製プラントの保守・点検作業において、ヘルメットにつけたカメラからの動画と身体加速度計からのデータを取得したという事例報告がある<sup>14)</sup>。それらのセ

ンシングデータから注視度合いや姿勢を判定し、操業者の行動について個体間の差異や個々の平均的行動と特定場面における差異を抽出している。

このような行動分析は、センシングデータの直接的な分析よりも、行動の行動素への分解や行動コーパスを介在した分析に移行している<sup>5),6),7)</sup>。行動素とは、身体の各部位の動きをクラスタリングすることで得られる身体動作の基本要素である<sup>8)</sup>。行動コーパスとは、マルチメディアコーパス<sup>9)</sup>から発展し、文章コーパスのように典型的な行動を集めたデータベースのようなものである<sup>9)</sup>。

センシングによって取得したデータを、行動素や行動コーパスに照らし合わせて、関連すると考えられる行動を抽出する、あるいは予測するというものである。なお、多くの場合、行動分析のためのセンシングは、音声データ、視線データ、移動データのような異なるモードのデータを取得対象とするマルチモーダルセンシングによって行われている<sup>10)</sup>。

高橋らは、イベント会場での人の動きに注目し、マルチモーダルセンシングによって音声、移動、顔の向きなどのデータを取得し、特徴的な動きを行動コーパスとしてまとめた<sup>9)</sup>。また、行動コーパスの利用までを考え、データ取得と同時にデータの記述性の向上を考慮し、タグの設計を行っている。彼らの行動コーパスは、大規模な動画像に対して、人の行動を検索キーとするシーン検索を目的に構築されたものである。一方、文献1),2),3)の手法を考えると、高橋らの方法を非熟練者と熟練者の行動の比較等に適用すれば知識抽出につながることは想像に難くない。

以上述べたように、ワークフローと行動分析の組み合わせは、人間の判断を含む行動から、慣習的な行動や、気づきの顕在化にとって有効な手段になると考えられる。この考えにもとづく研究例もある。例えば、竹林のグループは、行動に対するセンシングとワークフローを結び付けている<sup>11),12),13)</sup>。

彼らの注目した対象は、船舶ブリッジにおける音声コミュニケーションである。音声コミュニケーションを分析することによってブリッジ作業の効率化を目指している。音声データにもとづく行動記述を行い、音声行動コーパスを構築している。その際に音声データとワークフローを結び付けている。自然な形で生じる行動からのデータをワークフローに結び付けたことが重要である。特に、文献13)では、

\*<sup>†</sup> 東京大学大学院工学系研究科環境海洋工学専攻  
Department of Environmental and Ocean Engineering, School of Engineering, The University of Tokyo

<sup>††</sup> 東京大学人工物工学センター  
Research into Artifacts, Center for Engineering, The University of Tokyo

<sup>†††</sup> 東京大学新領域創成科学研究科人間環境学専攻  
Department of Human and Engineered Environmental Studies, Graduate School of Frontier Sciences, The University of Tokyo

- ・同じ業務ラベルが付与された一連の発話を取得する
- ・各発話に付与された個別行動ラベルの並びをワークフローと照らし合わせる
- ・ワークフローから逸脱した個別行動ラベルが付与されている箇所を抽出する

によって、実際にコミュニケーションエラーを抽出している。これはヒューマンエラー検知の例ではあるが、エラー発生箇所を分析することで、それを起こさないという施策に利用できる。しかし、彼らの方法には、主として二つの課題がある。

一つは、ワークフローと関連付けているが、音声コーパスの構築が目的になっており、ワークフローの修正、あるいは詳細化など、ワークフロー自体の変更という観点がない。もう一つは、個別行動ラベル付きデータへのワークフローの対応付けの仕組みが示されていない。目視による方法であり、自動化は難しいものと考えられる。

一般には、ワークフローによって、すべてのながれを記述することは難しく、記述されたものでも適応的に修正されることが多い。ワークフローを利用するためには、作成法や修正法が問題になる。これまでの研究では、観察、およびインタビューとヒアリングによって人間の判断を含む行動を分析し、ワークフローの作成や修正が行われてきた。最終的にはそうせざるを得ない部分もあるが、ある程度は自動的に分析されることが望ましい。

我々は、操業空間において、視線計測と位置計測を利用して操業者の行動をセンシングし、その結果をヒアリング等で得られる操業者の初期ワークフローに対応付け、ワークフローを詳細化し、その差分から操業者の知識として記述されていない慣習的な行動や気づきを顕在化するという手法を提案している<sup>15)</sup>。この際に、センシングデータから行動素・行動・行動群の抽出手法と、抽出されたものをいかにワークフローに対応付けるかが課題になった。これは、上述した竹林らの手法における課題と同様のものである。本稿では、ワークフローから抽出される行動素・行動・行動群と、センシングデータからのものを対応付けるという方法によってこの問題の解決を試みる。

## 2. 提案手法

### 2.1 ワークフローの詳細化手法

はじめに、先行研究<sup>15)</sup>におけるワークフロー詳細化手法の概略を示す。

- (1) ヒアリング等による初期ワークフローの作成
- (2) 視線計測と位置計測を利用した行動データの記録
- (3) BPMNと擬人化による初期ワークフローの再構成
- (4) 行動抽出と初期ワークフローへのマッピング
- (5) (4)の分析による、慣習的な行動や気づきの顕在化
- (6) ワークフローの詳細化

というものである。以下に(1)から(3)について簡単に示しておく。

#### 2.1.1 イベント抽出と初期ワークフローの作成

前もってわかっている作業プロセスに対して、操業者へのヒアリング等によって、“モニターの確認”、“キーボードの操作”、“場所の移動”、“警告音の発生”などの動作やイベントの抽出を行う。このイベントを識別子 $\tau$ で示す。次に、作業プロセスとイベントに基づき、動作とイベントの抽出と同様にヒアリングや操業現場の観察等によって、ワークフローを作成する。ここで、ワークフローの識別子を $W$ で示す。

### 2.1.2 行動データの取得とデータの同期

人間とシステムの相互間で生じるデータを操業者の行動データとする。本研究では、操業者の視線データと位置データ、および操業支援システムなどの操作履歴を行動データと考える。

本稿では、視線データを、EMR-8B という視線計測装置で取得する。操業者 $i$ の視線データを $i_i(n, t_n, x_m, y_n)$ とする。ここで、 $n$ は注視点の番号、 $t_n$ は記録開始からの時間、 $(x_m, y_n)$ は視野映像内の注視点の座標である。

位置データは、USBカメラで取得した映像を画像処理によって求める。予備実験の結果からは、環境光やノイズの影響でカメラ映像からの画像処理では正確に位置データが取得できないことがわかっている。

そこで、操業空間を、図1のように、表示装置(領域)や区画などの構成要素で表現することにした。この中で操業者の位置を領域 $R$ で表す。操業者の位置に関するデータは、 $i_p(\tau^i, T_\tau, R_\tau)$ とする。 $\tau$ はあるイベントの識別子、 $T_\tau$ はその発生時刻、 $R_\tau$ はそのときの操業者の位置である。

また、 $t_n$ と $T_\tau$ 、 $(x_m, y_n)$ と $R_\tau$ を関連付ける。操業支援システム等から生じるイベント(例えば、警告音の発生)に関しては、位置データに付与され、イベントの識別子と発生時刻を、その影響を受ける操業者に与える。

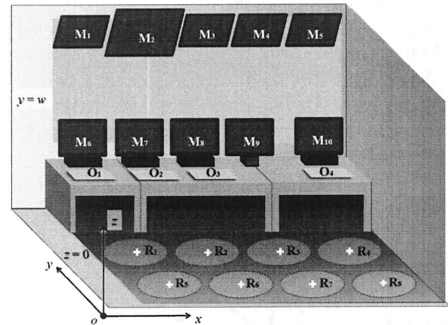


図1 操業空間の例

### 2.1.3 システムと空間の擬人化

提案手法では、受動的な行動も行動の対象とし、モニター、操作対象となるシステム、さらには、操業空間の場所を擬人化して考える。例えば、操業者 $x$ がモニターAの情報を確認した場合、モニターAは、“操業者 $x$ に確認された(見られた)”となる。

### 2.2 行動群(シーン)の抽出法

行動を認識する方法にはいくつかのものがあるが、本研究では、行動素から行動、行動から行動群という2段階で行動を扱う。

#### 2.2.1 行動素の抽出

植浦ら<sup>9)</sup>と同様に身体の各部位(目や頭部)の動きをクラスタリングすることで得られる身体動作の基本要素を行動素とする。例えば、“モニターの情報を確認する”という行動の際の視線データや位置データに対し、斜め右下から中央、中央から上部、上部から左上隅のように視線が移動すれば、それぞれが行動素に対応した結果とみなされる。このような行動素を、植浦らの方法のように操業者の目の動きなどを分析することで抽出する。本稿では、視線計測装置からの瞳の画像と位置データ取得のためのUSBカメラからの画像を利用する。瞳の画像は相

対的なものなので絶対的な位置が必要になるためである。

行動素は元の行動と対応付けられるので、行動素をまとめることによって行動を予測できるようになる。しかしながら、操業の現場においては、“モニターの情報を確認する”という一般化された行動ではなく、“モニターAの情報を確認する”というように決められた対象に対する行動になることが多い。また、作業全体では、行動素は分析には詳細すぎるものと考えられる。

そこで、行動素を以下のように扱う。

説明のため、モニターA, B, Cで構成される操業空間における操業者xの作業を考える。ワークフロー $W_1$ を、

$W_1$ : モニター群を確認→モニターAを確認  
→モニターBとCの比較→モニターBを確認

とする。図2左図にこのワークフローを、右図に擬人化後のワークフローをBPMNで記述した例を示す。

図3に、2.1.2項で述べた方法によって取得したデータを対象毎にガントチャートの表す。図3の横軸は時間である。例えば、モニターAの領域に注視点が存在し、 $r_b, r_{t+1}, \dots, r_{t+m}$ のように連続した場合、 $[r_b, r_{t+m}]$ を1つの区間として表示している。また、これを“区間”と呼ぶことにする。

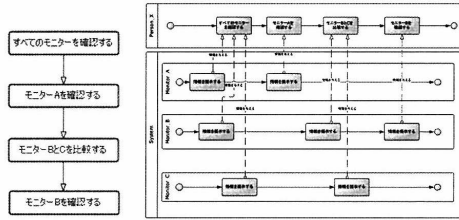


図2 初期ワークフローと擬人化による再構成

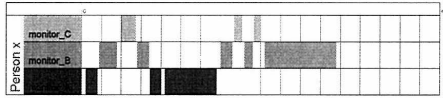


図3 モニターからみた行動

図3で示すデータはワークフロー $W_1$ に関連付けられている。図3の一つの区間は、何かしらの行動に対応する。行動素は、図3の一つの区間、あるいは複数の区間に、複数個存在する。この“区間”を、行動素をまとめる単位とし、行動素を階層的に管理する。

行動そのものではなく、行動の構成要素として行動素を考える理由は以下である。

センシングデータから得られる図3で示すような区間は、モニターAというような対象によって生じるものなので、一般化される知識を抽出するためには、抽象度の高い行動素で表すことが望ましい。

知識の詳細化には、区間内でのより詳細な行動パターンを調べる必要が生じるが、その際に行動素が利用できる。

次項で述べる行動群の抽出の際に、どこからどこまでを一連の行動とみなすかを機械的に処理することは容易ではなく、区間と区間の関係性を調べる必要が生じる。その際に、行動素が利用できる。

### 2.2.2 行動群の抽出

操業においては、いくつかの行動がまとまって意味のある行動になる場合も少なくない。そこで、いくつかの

行動をまとめることを考える。

行動をまとめるために、高橋らのインタラクションの統合手法<sup>9)</sup>を利用する。高橋らの方法では、人の様々なインタラクションに対してのシーン抽出であったが、提案手法では行動群の抽出と考える。この際に、擬人化と図3で示した区間を用いる。

例えば、ワークフロー $W_1$ を考える。この場合は、モニターと操業者がオブジェクトになり、1人の操業者xと3つのモニターA, B, Cのインタラクションを考えることになる。本稿では、高橋らと同様の方法で、オブジェクトをノード、インタラクションをリンクとしたグラフの時間変化を調べる。

操業者xがモニターAを見る場合、モニターAが操業者xに情報を与えることになる。この場合は、Aからxに有向リンクが生じる。操業者xがモニターAに何かしらの情報を表示させる場合、xからAに有向リンクが生じる。操業者xがモニターAをみて、何かしらの作業を行うという場合は、xとAに双方向のリンク、すなわち無向リンクが生じることになる。

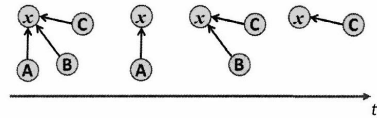


図4 インタラクションのグラフ表現

ワークフロー $W_1$ に対するものが図4である。センシングデータから、この図を求めるためには、分割の任意性を解消する必要がある。本稿では、高橋らの方法<sup>9)</sup>に従い、リンクの強度、コミュニティの強度、ノードの貢献度を求め、自動的に分割する。

$L_{ij}(t)$ をノードiからノードjに向かう時刻tにおけるリンクの強度とし、

$$L_{ij}(t) = (1/w_1) I_{ij}(t, t+w_1) \quad (1)$$

で与える。ここで、 $I_{ij}(t, t+w_1)$ は、時刻tから時刻の $t+w_1$ 間のノードiからノードjへのインタラクションの総時間である。 $w_1$ はパラメータである。

いくつかの行動をコミュニティとしてまとめる。あるコミュニティをCとする。Cはいくつかのノードで構成される。コミュニティに存在するリンクの強度によって、時刻tのコミュニティの結合力 $F_C(t)$ を求める。

$$F_C(t) = 1/|C|^2 \sum_{ij} T_C(t)[i,j] \quad (2)$$

ここで、iとjはCに属するノードであり、|C|はコミュニティに属するノードの数を示す。また、

$$T_C(t)[i,j] = \begin{cases} (1/w_2) \cdot I_{ij}(t-w_2, t) & (t-t_s > w_2) \\ (1/(t-t_s)) \cdot I_{ij}(t_s, t) & (t-t_s \leq w_2) \end{cases}$$

である。 $t_s$ はコミュニティCが生成された時刻、 $w_2$ はパラメータである。

コミュニティを構成するノードに対して、ノードのコミュニティへの貢献度を求める。ノードkの貢献度 $P_C(k, t)$ を、

$$P_C(k, t) = 1/|C|^2 \sum_{ij} A_C(k, t)[i,j] \quad (3)$$

とする。iおよびjがkと一致しない場合は、 $A_C(k, t) = 0$ とし、そうでない場合、

$$A_C(k, t)[i,j] = \begin{cases} (1/w_3) \cdot I_{ij}(t-w_3, t) & (t-t_s > w_3) \\ (1/(t-t_s)) \cdot I_{ij}(t_s, t) & (t-t_s \leq w_3) \end{cases}$$

とする。ここで、 $w_3$ はパラメータである。

高橋らは、リンクの強度、コミュニティの強度、ノードの貢献度によって、コミュニティの変化を抽出した<sup>9)</sup>。本稿でもその方法に従う。

(a) 新たなコミュニティの生成

$$L_{ij}(t) + L_{ji}(t) > c_1 \quad (4)$$

(b) コミュニティの拡大

$$L_{ij}(t) + L_{ji}(t) > c_2 F_C(t) \quad (5)$$

(c) コミュニティ構成の変化

$$L_{ij}(t) + L_{ji}(t) > c_3 (F_{C1}(t) + F_{C2}(t)) \quad (6)$$

(d) コミュニティの縮小

$$P_C(k, t) < c_4 F_C(t) \quad (7)$$

(e) コミュニティの消滅

$$F_C(t) < c_5 \quad (8)$$

ここで、 $c_1, c_2, c_3, c_4, c_5$  はパラメータである。

以上の方法で一連の作業が動的グラフとして抽出される。また、抽出した行動素、行動、行動群を対応付けて管理し、行動コーパスとして蓄積する。

### 2.3 ワークフローへの対応付け

前節で述べた方法によって求められた動的グラフを初期ワークフローに結び付ける。

はじめに、図2のようにBPMNで表されたワークフローから、リファレンスとなるグラフを抽出する。工程に対して適当に時間を設定し、前節の方法によって動的グラフを抽出すればよいが、精度の観点から、本稿では手作業によって動的グラフに変換する。例えば、 $W_1$ は図4となる。

次に、図3のようにガントチャートのまとめられたセンシングデータから前節で述べた方法によって動的グラフを抽出する。

2つの動的グラフをグラフマッチングによって対応付ける。

### 3. 実験結果と考察

提案手法の有効性を確認するために、先行研究<sup>10)</sup>と同じ実験環境で、操業者の行動を分析する。初期ワークフローとして、

$W_2$ : モニター1を確認→すべてのモニターを確認→モニター3を確認→モニター9を確認→モニター6を確認→モニター4を確認→モニター6とモニター4を比較→モニター7を確認→モニター8を確認→モニター1を確認

を与える。

図5にBPMNで記した初期ワークフローと対応する動的グラフを示す。

二人の被験者に対して、三つの方法で指示を出すことによって、工程を進めてもらった。1回あたりの実験時間は約100秒であった。得られた視野画像と視線データは約3000である。各画像を目視によって調べ、各々の工程作業の開始時間と終了時間を求め、ガントチャートの表現したものが図6である(一例を示している)。これから(1)から(8)式によって動的グラフを抽出し、図5下と比較する。詳細は講演時に述べる。

### 4. まとめ

本稿では、計測によって得られた行動素のワークフローへの自動対応付け手法を検討した。高橋らの方法を利用して、計測データから動的グラフを生成し、初期ワークフローから得られる動的グラフとの比較(グラフマッチング)により、半自動的な対応付けが可能になることを示唆した。

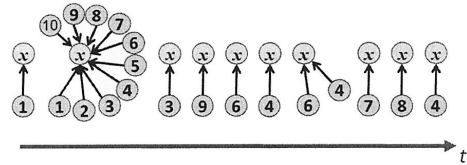
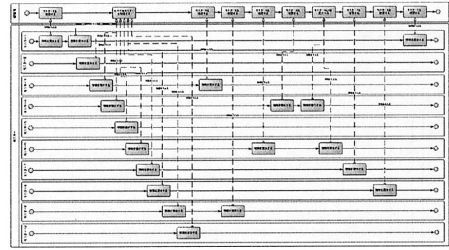


図5 初期ワークフロー $W_1$ と動的グラフ表現

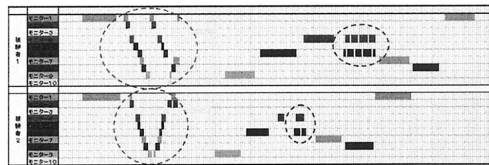


図6 計測データのガントチャートの表現

### 参考文献

- 1) Harper, D.: Working Knowledge: Skill and Community in a Small Shop, Univ. of California Press, 1987.
- 2) Kimura, F., Ariyoshi, H., Naruko, Y., Yamamoto, H.: Capturing Expert Knowledge for Supporting Design and Manufacturing of Injection Molds, Annals of the CIRP, 53.1, pp.147-150, 2004.
- 3) 成子由則: モノづくりに関する知識・ノウハウの伝承, 情報管理, 49.8, pp.439-448, 2006.
- 4) Hiekata, K., Yamato, H., Oishi, W., Sasaki, Y., Sato, K.: A Knowledge Management Framework for Marine Propeller Design, Proceedings of Intl. Conf. on Computer Applications in Shipbuilding 2007, pp. 79-89, 2007.
- 5) 植浦総一郎, 岩井儀雄, 谷内田正彦: 准教師有リクラスターリングによる行動素抽出, 情報処理学会研究報告-コンピュータビジョンとイメージメディア, 2008.36, pp.29-36, 2008.
- 6) Takahashi, M., Ito, S., Sumi, Y., Tsuchikawa, M., Kogure, K., Mase, K. and Nishida, T.: A Layered Interpretation of Human Interaction Captured by Ubiquitous Sensors, Proc. of 1st ACM CARPE Workshop, pp.32-38, 2004.
- 7) 角康之, 伊藤慎宣, 松口哲也, Sidney Fels, 間瀬健二: 協調的なインタラクションの記録と解釈, 情報処理学会論文誌, 44.11, pp.2628-2637, 2003.
- 8) 武田一哉, 河口信夫, アブットフセイン, 板倉文忠: 音声・映像・運転行動コーパスを用いた車内情報インタフェースの研究, FIT2003, 2003.
- 9) 高橋昌史, 角康之, 伊藤慎宣, 間瀬健二, 小暮潔, 西田豊明: 時系列イベント発見のためのグラフクラスターリング手法の提案, 情報処理学会論文誌, 49.6, pp.1942-1953, 2008.
- 10) 石塚満: マルチモーダル擬人化インタフェースとその感性基盤機能, 平成11年度~平成15年度 日本学術振興会 未来開拓学術研究推進事業研究成果報告書, 2004年3月.
- 11) 青島大悟, 鈴木敦志, 桐山伸也, 杉山岳弘, 竹林洋一: マルチモーダルセンシングによるブリッジ業務分析システムの開発, 第22回人工知能学会全国大会講演論文集, CD-ROM, 3D3-08, 2008.
- 12) 鈴木敦志, 望月研吾, 桐山伸也, 竹林洋一: ブリッジ業務支援のための音声行動コーパス, 日本音響学会 2008年春季研究発表会, 2-4-15, 2008.
- 13) 桐山伸也, 鈴木敦志, 青島大悟, 本間永愛, 竹林洋一: 安全航行のための船員の音声コミュニケーション分析, 日本音響学会 2008年秋季研究発表会, 1-1-22, pp.53-54, 2008.
- 14) 人間生活工学センター, 東洋エンジニアリング: 石油プラント保守・点検作業支援システムの開発, H19年度プロジェクト事後評価用資料, 2007.
- 15) 樺天甲, 白山晋, 榊方和夫, 大和裕幸: 操業者の行動分析にもとづくワークフロー詳細化手法, 情報処理学会研究報告(MP S), Vol.2008, No.126, pp.21-24, 2008.