

野球番組ダイジェストのためのアナウンス文の自動生成

4D-4

畑田のぶ子\* 金 淵培\* 浦谷則好\*\* 山田一郎\*

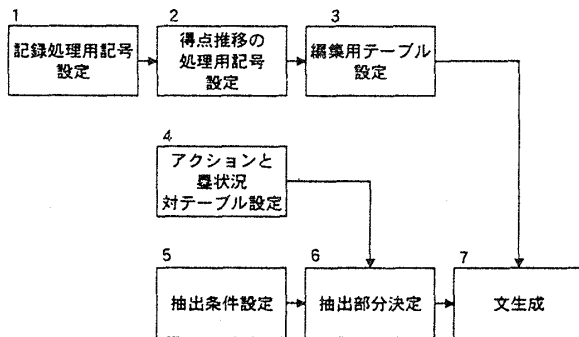
(\*NHK 放送技術研究所 \*\*ATR 音声翻訳通信研究所)

1.はじめに

いろいろなイベント（スポーツ、音楽、囲碁、将棋）の内容を表現する方法に、記号を用いる方法がある。特にスポーツの中で、野球は内容が構造化され、状態が明確に定義できるのでこの方法が用いられる。この記録用の記号（スコア表）を利用して、状態やその時起こったアクションをアナウンス文に生成し、映像とリンクすることにより、ダイジェスト番組の自動生成を行うことを試みる。今回は、ダイジェスト番組の自動生成のため、野球のスコア表の記号を用い、ダイジェスト番組用のアナウンス文の自動生成を行った結果を報告する。

2.アナウンス文生成のシステム構成

文生成システムのブロック図を1図に示す。



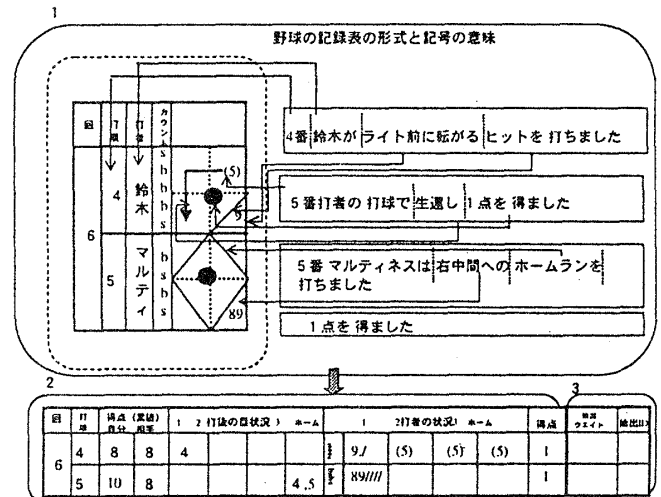
1図 文生成システムのブロック図

1～3までは、記録表の記号や状態を、計算機処理用の記号に変換し、キーワードとテンプレート文を付加し、条件抽出や文生成の準備をするステップである。

4は実際のゲーム展開が記述してある記録表を処理用のデータ構造に変換する部分である。

7は文の生成を行う部分である。

野球の記録表の形式と記号の意味を2図-1に示す。野球の記録表は攻撃側に焦点をあてて記述しており、打者が単位となり、「打者が打席に入り、何かのアクションが起こり、ランナーとなり（あるいはランナーにならず）、アウトになるか、または生還し1得点を得る。」で表わされる。2図-1で示す1つのボックスに、これら状態が記号で記録され、先頭打者から3アウトになるまでが、1攻撃回となる。



2図 野球の記録表とアクション・壘状況テーブル

1) 記録処理用記号の設定

試合の記録用の記号を、A) 記号の動作主や目的格が明らかで、特に明示しなくてよいもの（盗塁、ヒット、三振）、B) 動作主や目的格、アクションの方向、目的格の質などを明示する必要のあるもの（送球の経路、エラー）の2種類に分け、各記号に対し、キーワード、面白度ウエイト、テンプレート文を付加したテーブルを作成する。

2) 得点推移の記述

面白い状態を抽出するため、得点推移（逆転、同点）、得点の有無、特別な条件（9回裏、満塁）、キーワード、文テンプレート、面白度のウエイトを対にしたテーブルを作成する。

これら1)、2)のキーワードと面白度ウエイトは、抽出条件の指定に、テンプレート文は文生成時

に使用される。

3) 編集用テーブル設定

その他に、つぎのテンプレート文を用意する。

- ・対戦チーム名や回の開始、終了時の得点など、状況説明や見出しにあたる文
- ・試合の途中部分をまとめあげる文
- ・守備位置のテーブル、メンバーテーブル、打球の方向テーブル、打球の質テーブル

4) アクションと塁状況対のテーブル設定

試合の記録表を、1) で定義された記号を用い、2 図-2で示すような「アクション」と「塁の状況」の対のテーブルに変換する。このテーブルは

- ・ある打者の状態遷移 (2 図-1の1ボックス)
- ・その打者によって引き起こされた塁の状況
- ・時系列を示す情報 (回、表裏)
- ・両チームのその時点での得点合計
- ・抽出用のウエイト、抽出IDエリア (2 図-3)

を合わせて1つの単位としたテーブルである。

5) 抽出条件設定手段

- ・抽出条件設定の準備

抽出条件は、キーワード (1と2のテーブルで設定されたもの) と生成文の長さ、ユーザーの好みのチームや選手が指定できる。ユーザは、キーワードの面白度ウエイトを指定時に自由に変更できる。

- ・指定の方法にはつぎの二通りがある。

- A) 条件を逐次指定する方法
- B) すでに準備してあるものを使う方法

この二種類の方法と、上記の抽出条件の選択、ウエイトの操作により、ユーザは、ニーズに合った抽出条件を作成することができる。

表1 ダイジェスト用の抽出条件テーブル

キーワード	ユーザー面白度
逆転さよなら勝ち	10
さよなら勝ち	9.5
逆転	9
同点	8
追い上げ	7
ホームラン	6
満塁 (得点)	5
先制点	4
得点	4
満塁 (得点なし)	3
盗塁	2
犠牲バンド	2
犠牲フライ	2
三塁打	1
二塁打	1
一塁打	1

一般的なダイジェスト用の抽出条件テーブルを表1に示す。

6) 抽出部分の決定

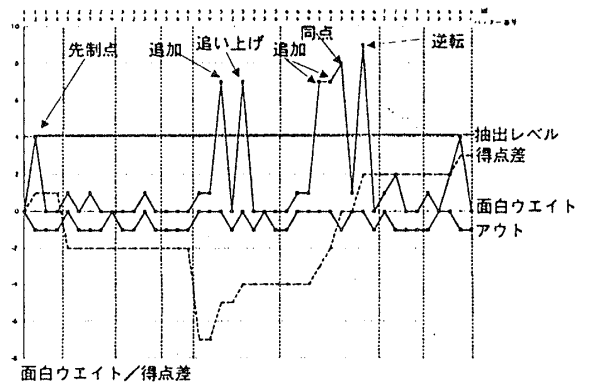
抽出条件を用い、「アクションと塁状況対のテーブル」の各ラインの面白度ウエイトを計算する。抽出のための最小のウエイト (以後、レベルという) を決め、そのレベル以上を抽出部分とし、レベル未満をまとめあげ部分とする。

このレベルは下記の条件により上下させる。

- ・抽出ポイントが多すぎるばあい、(レベルを上げて抽出ポイントを少なくする)
- ・抽出ポイントが少なすぎる場合 (レベルを下げて抽出ポイントを多くする)
- ・短いダイジェストがほしい (レベルを上げて抽出ポイントを少なくする)
- ・詳しいダイジェストがほしい (レベルを下げて抽出ポイントを多くする)
- ・レベルを零に設定することにより、試合経過全体を抽出できる。
- ・レベルを最高得点以上にすることにより、結果のみを抽出できる。

得点とレベルの概念図を3図に示す。

レベルのコントロールをおこない、抽出部分を決定し、抽出点にID (文番号) を付加する。



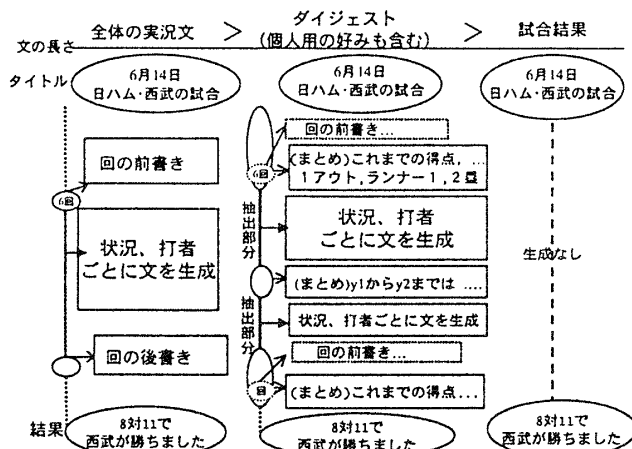
3 図 面白度ウエイトと抽出レベルの関係

7) 文生成

文生成で生成される文の種類を、時系列を軸として、抽出部分とまとめ上げ部分の関係で示すと4図のようになる。

4 図で示されるように、「全体の試合経過の説明文」から、「ダイジェスト (好みの抽出部分に焦点をあてた説明文も含む)」を経過して「試合結果のみの説明文」に至る過程は、抽出部分が試合経過全体から、段階的に少なくなり、抽出部分がなくなっ

たと考えることができる。



4図 文の種類

6) の処理がなされた ID 付きのアクションと塁の状況対のテーブルと、文生成のためのテンプレートを  
用い、文を自動生成する。全体の試合経過の説明文の生成結果の例を 5 図に、ダイジェスト文の生成結果の例を 6 図に示す。

6月14日、西武球場、日本ハムファイターズ対西武ライオンズの試合です。

....

1 回の裏、西武ライオンズの攻撃  
一番大友はフォーボールで一塁に進みました。  
2 番松井はサード左を抜く 2 塁打を打ちました。  
この間に、一番ランナーはホームインしました。  
3 番高木は、セカンドゴロをファーストに送球され、  
1 アウトになりました。  
この間に、2 番ランナーは 3 塁へ進みました。  
4 番鈴木健はセンターフライで 2 アウトになりました。  
松井は 4 番の打球をセンターからキャッチャーへ送球され 3 アウトになりました。  
この回の得点は 1 点となりました。

2 回の表、日本ハムファイターズの攻撃  
これまでの得点は、日本ハムファイターズが 0 点、西武ライオンズが 1 点です。

.....

5 図 全体の試合経過の説明生成結果

6月14日、西武球場、日本ハムファイターズ対西武ライオンズの試合です。...

1 回の裏、西武ライオンズの攻撃、ノーアウト、ランナー 1 塁 2 番松井のサードの左を抜く 2 塁打で 1 点

を取り先制します。

5 回の表、日本ハムファイターズの攻撃、ワンアウト、ランナー 2、3 塁

8 番上田のレフトの前へのヒットで 2 点を取り 2 対 1 で逆転しました。

さらに、ツーアウト、ランナー 2 塁

1 番田中のライトの前へのヒットで 1 点を取り 3 対 1 と得点を追加します。...

これまでの得点は、日本ハムファイターズが 8 点、西武ライオンズが 4 点です。

6 回の裏、西武ライオンズの攻撃、ワンアウト、ランナー 2,3 塁、

1 番大友のファーストの上を抜くヒットで 1 点を取り 5 対 8 と詰め寄ります。...

さらに、ワンアウト、ランナー 1、3 塁

3 番高木大のライトの右を抜く 2 塁打で 2 点を取り 8 対 8 の同点としました。

5 番マルチネスのセンターライトの間へのホームランで 2 点を取り

10 対 8 で逆転しました。...

この結果、日本ハムファイターズの得点は 8 点、西武ライオンズの得点は 11 点で西武ライオンズが勝ちました。

6 図 ダイジェスト文生成結果

### 3. 今後の課題

映像込みのダイジェスト番組生成には映像とリンクしたスコアデータの効率的入力手法が必要になる。これに関しては市販のソフトやデータを利用することが考えられる。野球以外のスポーツへの適用に関しては、現在検討中である。より高度なダイジェストの解説文の生成のためには、現在のような固定された表現でなく、生成のための語彙を増やし、「状況の変化をみながらシナリオを自動生成し、適切な語を選択する」ような知的な処理が必要になる。これについても、今後検討する予定である。

### 【参考文献】

畑田のお子、相沢輝昭 「記録表からの文生成の一手法」、情報処理学会第 23 回全国大会講演論文集 1981.