

身体的コミュニケーションを活用した対話システム「おしゃべり 6256」の開発

筑波技術大学 産業技術学部 産業情報学科

中山 睦 内藤一郎

要旨:コミュニケーションにおいて、表情や顔つきといった身体的な表現は重要な役割を担っている。我々は、対話システムにおけるこの身体的コミュニケーションの有効性について検討を行なっている。提案する対話システムは、人間側の語り掛けに対し文字や音声などによる応答だけでなく、実装した感情モデルに基づいて画像やロボットなどに身体的な動作を行なわせることでより真に人間らしいコミュニケーションを行なうものである。今回、入力文に対する反応にCGアニメーションやロボット等による身体的な動作を加えて応答する対話システム「おしゃべり 6256」を試作し、評価実験を行った。本稿では、実施した結果を報告するとともに、その結果を基に将来的な取組みの可能性についても考察する。

キーワード: 身体的コミュニケーション, 感情モデル, 対話システム

1. はじめに

我々は普段の会話の中に単に言葉による情報だけでなく、うなずきや手ぶりなどの身体動作や表情といった、言葉によらない情報も取り入れることで、より深く、円滑にコミュニケーションを成り立たせる。これによって場の盛り上がりや相手との一体感を感じ、人との関わりを実感させている。このメカニズムを機器やシステムに導入し、人間側の語り掛けに対し、文字や音声などによる応答だけでなく、画像やロボットなどの外部機器に身体的な動作を行なわせることでより真に人間らしいコミュニケーションシステムを実現できるものと思われる。そこで、入力文に対して応答する対話システムに、身体的コミュニケーションを表現するためのCGアニメーションやハンドパペット型ロボット、LEDライトを装備したシステムを開発し、対話システムにおける身体的コミュニケーションの有効性を検討することにした。身体的コミュニケーションの応用イメージを図1に示す。



図1 身体的コミュニケーションの応用イメージ

2. 対話システム「おしゃべり 6256」

2.1 システム構成

今回開発した対話システムは入力文に対して応答文を

作成する対話エンジン、応答文をCGアニメーションとともに表示する表示エンジン、ハンドパペット型ロボットとLEDライトで身体的コミュニケーションを表現する制御エンジンから構成した。

各エンジンの特徴は次の通りである。

○ 対話エンジン

- ・人工無能型による対話システムで自然な会話を実現
- ・入力文に合わせて顔文字アニメーションを表示
- ・学習機能を装備

○ 表示エンジン

- ・応答文を表示
- ・感情モデルによりCGアニメーションを変化

○ 制御エンジン

- ・ハンドパペット型ロボットで口（くち）パクや身振りなどの身体動作を表現
- ・LEDライトで感情を多様な色と点灯パターンで表現



図2 対話システム「おしゃべり 6256」

今回開発したシステムを図2に示す。システムの名称は、著者のニックネームを用いて「おしゃべり6256」(6256は、「むつごろう」と読む)とした。ハンドパペット型ロボットの外観(カバー)は、ハンドパペット「パペットケロロ軍曹」(バンダイ)を使用した。LEDライトのカバーは学研「大人の科学マガジン Vol.29 AKARI 折紙」[1]を参考して和紙を用いて製作した。

また、今回のシステムは聴覚障害者の利用を考え音声による入出力でなく、キーボードで入力し、モニタに応答文を表示するユーザーインターフェースを採用した。

「おしゃべり6256」では、ユーザーから入力があると最初に対話エンジンで応答文の作成とシステムの感情値(以降、キーワードの感情値と区別するため、ステータス値と記す)の決定がなされ、その内容をPC内部のTCP/IP通信(サーバー・クライアント)を利用して表示エンジンへ送る。表示エンジンでは、受け取ったステータス値に合わせたCGアニメーションを使って応答文を表示するとともに、ロボットの動作とLEDライトの点灯のパターンをステータス値に合わせて決定してシリアル通信を用いて制御エンジンに渡し、制御エンジンでロボットとLEDライトを制御する。各エンジンの通信の概要を図3に示す。



図3 各エンジンの通信の概要

2.2 対話エンジン

対話エンジンは、秋山の対話システム『ノビィ』[2]をベースに、Rubyを使用して開発した。Rubyはオブジェクト指向言語であり機能ごとにオブジェクトにまとめることでさまざまな変更も容易にできる。また、テキスト処理の機能が豊富であるため、この言語を選択した。また、今回のシステム開発においては、身体的コミュニケーションの有効性について検討することが目的であり、人間の知的活動の検討や会話内容を必要以上に充実させることが目的ではない。したがって、今回、対話システムはシステム構築が容易であり、会話をしているふりをし、人間側にあたかも自然な会話をしていると思わせることができる人工無能型を用いることにした。

対話エンジンでは、文章が入力されると文章を形態素解析し、その結果をもとに応答文を作成、後述の感情モデル

に基づいてステータス値が決定される。応答文を作成したあと、それまでに処理された内容をまとめて学習し、ファイルに追加保存して、次の応答文作成に活用する。なお、形態素解析には、奈良先端科学技術大学院大学で開発された形態素解析システム「茶釜」を使用した[3]。

対話エンジンを起動すると図4の対話ウィンドウが表示される。文章を入力し、続けてエンターキーを押すか「話す」ボタンを押すと、入力された文章はウィンドウ内の吹き出し部分に、入力文と応答文はログ表示の部分に表示される。さらに、ステータス値に合わせて顔文字アニメーションが変更表示される。

対話エンジンには、5種類の応答アルゴリズムと3種類の学習アルゴリズムが装備されている。

応答アルゴリズムは、以下に示す5種類を用意し、入力文に対して応答文を作成する。

WhatResponder

「～って何?」という聞き返しをする。(～の部分には入力文が入り、オウム返しをする)

RandomResponder

あらかじめ用意してある文字列(応答文)のうちランダムに選んだ1つの文字列を返す。

PatternResponder

入力文の中に用意してあるキーワードがあればパターン辞書の先頭からパターンマッチを行い、マッチした応答例の中から応答文を選択。

TemplateResponder

入力文から一定の言葉を抜き出し、穴埋め状の文章のテンプレートに埋めてそれを応答文として返す。(穴埋めは1語、2語。と用意)

MarkovResponder

形態素解析をした上でマルコフモデルに従い、新しく文章を生成しそれを応答文として返す。

入力文に対して、使用される応答アルゴリズムを乱数で選択し、応答文を作成するとともに、ステータス値を決定す

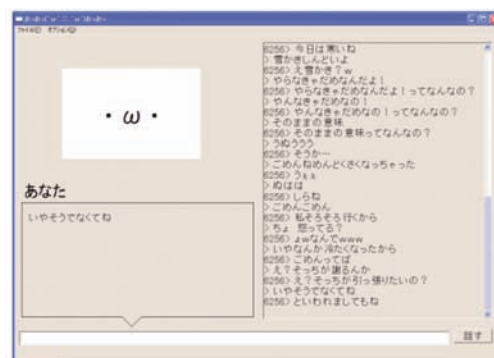


図4 対話ウィンドウ

る。各キーワードに感情値を持たせ、感情値に応じて1次元モデル(上機嫌-不機嫌)の感情表現を行う。入力文に感情値を持ったキーワードが存在する場合(例:「ばか」-5、「かわいい」+5など)には、その値に応じてステータス値が変化する。また、入力文に感情値を持ったキーワードが存在しない場合には、徐々に(0.5ずつ)平常値(ステータス値が0)に復帰するようにした。

通常、感情は「喜怒哀楽」の4つで表現されるため、2次元モデルとして感情を表現することも考えられるが、「喜」と「楽」の2つは感情表現が似ていて区別が難しく、また、座標軸の直交性にも疑問があるため、今回は1次元モデルの感情表現方法を採用した(図5)。

また、本システムでは、応答アルゴリズムによって応答文が作成されたあとにさらに語尾(顔文字を含む)の後付けを行っている。同じ応答文でも語尾が変わるだけで応答文の雰囲気は全く変わり、ユーザーを飽きさせない会話を実現できると考えた。後付けされる語尾は乱数により決定される。

応答文作成後、入力文の学習を行う。学習アルゴリズムは以下の3種類である。

入力文の学習

プログラムを終了するときに、それまでの入力文をランダム辞書のファイルに追加保存する。

キーワードの学習

形態素解析システム「茶釜」を利用して入力文を形態素に分解し、品詞を決定、そこからキーワードを抜き出し、感情値を追加してキーワードごとに応答例をパターン辞書のファイルに追加保存する。

マルコフモデルに従った学習

入力文を形態素解析し、形態素のつながりを記憶した上でプレフィックス(前にある形態素)とサフィックス(後ろにある形態素)のつながりを学習し、2プレフィックスを選択した上で1サフィックスを選択するといった処理を繰り返し暗記させ、より人間らしい会話を成立させる。会話で学習を重ねると応答文が多様になる。

2.3 表示エンジン

表示エンジンはProcessingで開発した。ProcessingはJavaをベースにした言語であり、オブジェクト指向言語であるとともに、通信機能、グラフィック機能が充実していることから選択した。表示エンジンには5パターンのCGアニメーションを用意し、ステータス値によってCGアニメーションを変更するよう設定した。吹き出し部分に対話エンジンから受け取った応答文を表示する。また、対話エンジンで決定したステータス値を表示エンジンで受け取った後、表示エンジン内でさらに乱数を使用して条件分けし、その中で

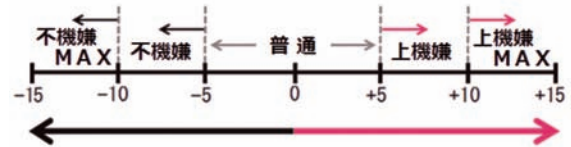
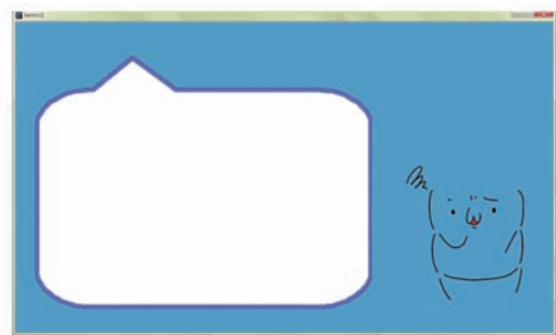


図5 1次元モデルの感情表現



(a) 感情「上機嫌 MAX」(ステータス値: 10 ~ 15)



(b) 感情「不機嫌」(ステータス値: -5 ~ -10)

図6 CGアニメーションの一例

poss 値(制御エンジンの動作パターン)を決定する。決定されたposs 値を、制御エンジンへシリアル通信で渡す。CGアニメーションの一例を図6に示す。

2.4 制御エンジン

制御エンジンはArduinoで開発した。ArduinoはC言語をベースにした言語で制御し、拡張ボードも多く、開発にはUSBケーブルとArduinoボードだけというシンプルな構成になっている。そして比較的簡単に制御が可能であることから選択した。ロボットとLEDライトの動作を同時に行うために、それぞれ別のArduinoで制御を行っている。ハンドバペット型ロボットは、9個のサーボモータをArduino MEGA2560で制御することで、ポーズや動作を行う。LEDライトは2つのフルカラーLEDを、Arduino UNOにより制御した。

本システムでは、応答文、CG、ロボット、LEDライトのそれぞれで感情表現を行っている。感情表現は、次の会話に重なることが生じないように、短く簡潔にするように心がけた。

制御エンジンでは、表示エンジンから受け取った possible 値に応じて 16 通りのロボットの動作ならびに LED ライトの点灯パターンを制御している。

2.5 身体的コミュニケーションの表現

CG 作成において、1 次元モデルの感情表現を「普通」、「上機嫌」、「不機嫌」と大きく 3 つに分けた。インターネット上で調査したところ喜怒哀楽のうち「喜」と「怒」だけに限定して見ると、感情を色で表した場合「喜」はピンク、「怒」は赤というイメージが強いことがわかった。この状態だと 1 次元モデルで表す際どちらも同色系になってしまい区別が難しくなる。だが、「喜」と「哀」だけに限定して再調査してみると、「喜」はピンク、「哀」は青という結果が得られ、区別がつきやすい色に分けられる上、違和感もないことがわかった。この結果をもとに「上機嫌」を暖色系、「不機嫌」を寒色系とすることに決めた。そしてその中間である「普通」には中間色に当たる緑色を採用した。普通の状態ならば中間色（緑）の状態にしておき、そこから上機嫌になるほど暖色系の色が濃くなり、不機嫌になるときほど寒色系の色が濃くなるように CG を作成した。また、背景色だけでなく、感情に合わせた CG アニメーションもつけることで、より感情が伝わりやすくなるを考え、CG アニメーションも付加した。アニメーションは上機嫌になるほど嬉しそうな動きに、不機嫌になるほど怒っているような動きになるように心がけて作成した（図 6）。

また、LED ライトの点灯パターンの検討に際して情報システムコース・コンピュータ科学コース 4 年生に意見をもらったところ、「喜」と「哀」に限定して結果を見ると先ほどあげた色の振り分け方と同じ意見が多く、信頼性も高いと判断して LED ライトにも採用した。色の表示時間に関しては「喜」のときはテンポが早め、「哀」のときはテンポが遅めという意見が多く、これも LED ライトを制御する際に参考にした。「上機嫌」の時には色を複数組み合わせる色での切り替えを行うようにしたが、「不機嫌」の時には単色のみで点滅を行うように差別化した。また例外として、「上機嫌」の数値が高い時には時間差で全色を表示するというレインボー色を出力するようにした。この色はユーザーが見る分にも楽しめると考慮し選択した。

ロボットの制御パターンについては、表情を出せないかわりに動きだけで上機嫌、不機嫌を区別し表現できるように考えて作成した。先ほどあげた色の表示時間に関して「喜」のときはテンポが早め、「哀」のときはテンポが遅めという意見を参考にし、「上機嫌」のときには機敏な動きを、「不機嫌」のときには鈍い動きをするように考慮した。「普通」のときには日常私たちが会話の中でも行っているうなずきや手ぶりを制御パターンの中に取り入れた。

3. 評価実験

本学の情報系 2 年生 19 人を対象とした評価実験を行った。3～4 人のグループ毎に、必ず全員が操作するように指示した上で今回製作した「おしゃべり 6256」を 5 分間体験してもらった。

体験後、アンケートを実施し、対話システムの評価を行った。アンケートの内容は、次の通りである。

設問① 対話システムの使用について

1. このシステムを面白いと思ったか?
2. 全体的に見やすかったか?
3. このシステムは使いやすかったか?
4. このシステムを今後も使ってみてみたいと思ったか?
5. 自分もこんなシステムを作ってみてみたいと思ったか?

設問② 感情表現の方法としての効果について

1. 会話の内容
2. CG（別のディスプレイに表示したウィンドウ）
3. ロボットの動作
4. LED ライトの色
5. 感情の表現の仕方は適切だったか?

これらの質問に対し 5 段階（悪い-普通-良い）で評価してもらった。また、その他に自由記述の欄も設けた。

4. 評価結果と考察

アンケートの回答結果を図 7、図 8 に示す。

「面白い」の評価は平均値が 4.3 であり、「見やすい」、「使いやすい」、「今後も使ってみてみたい」の評価に対して両側検定、有意水準 1% で高い評価が得られた。標準偏差も他の項目に比べて小さく、本システムを経験してもらった学生（以降、被験者と記す）の多くが高く評価していたことがわかる。

また、グラフから見てわかるように、操作感は必ずしも高くなく、平均値も 3.0 で「どちらでもない」となっている。一般に操作性の悪いシステムでは、「面白い」などのシステムの総合評価が低くなる場合が多いが、本システムでは「面白い」の評価が有意に高い。このことから、このシステムにはユーザーに楽しさや癒しを与える効果があると思われる。

感情表現の方法の中で一番評価が高かった「CG」の評価の平均値は 3.9 であり、「会話の内容」の評価に対して両側検定、有意水準 5% で高い評価を得た。CG の評価が応答文に対して有意に高く、自由記述でも「CG の表現の仕方がわかりやすかった」とコメントがあった。このことから当初考えた CG の背景色の振り分け方、アニメーションの表現方法は適切で効果があったと思われる。また、今回のシステムで対話エンジンに付け加えたのは「感情表現」

であり、会話にも感情表現が含まれているのに関わらず、会話に比べてCGが有意に高い結果になったことから、被験者が会話の内容よりもCGの変化に面白さを感じていたことを示していると考えられ、計画段階で考えた身体的コミュニケーションが面白さに結びついているのだと思われる。このことから、本システムの身体的コミュニケーションには効果があると言える。

一方、感情表現において当初期待していた、ロボットの動作やLEDライトの点灯の評価は有意差こそなかったが、CGと比べて低くなった。CGは、早い段階で実装ができたため、開発期間が特に長く(2ヶ月以上)さまざまな意見やアドバイスを積極的に取り入れて作成できたのに対し、ロボットやLEDライトは完成時期が遅れ、開発期間は1~2週間ほどの短い間となってしまった。そのため、完成度がCGに比べて低く、不十分などが見られる。ロボットの動作にはまだ少々きこえない点があったのでその部分が滑らかになるよう作り直すことは効果的であると思われる。

また、CGは入力したあと応答文と同時に変更、出力ができるためタイムロスもあまりなかった。それに対し、ロボットやLEDライトはさらにシリアル通信を経て制御を行っているため短時間ではあるがタイムロスが生じた。このことは、被験者からも自由記述で指摘されている。この通信の遅れを解消し、ロボットの動作やLEDライトの点灯の完成度をあげられるとさらに良い評価が得られ、もっと面白いと思ってもらえるのではないかと考えている。

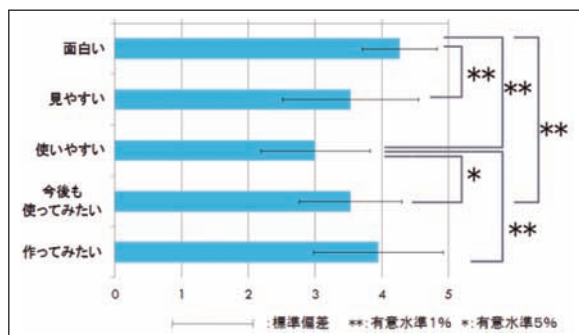


図7 設問①の評価結果

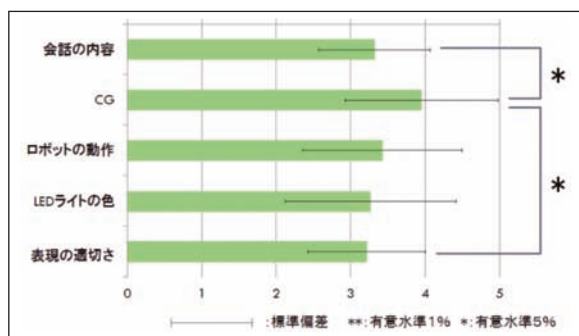


図8 設問②の評価結果

以上のことから、ロボットやLEDライトは、会話の腰を折らないようにと短く、完結な表現を心がけたが、表現方法にはさらに改善が必要だとわかった。ただし、各表出方法には有意差がなく、自由記述でリアクションやライトの色などが良かったという意見も出されていることから、よりわかりやすい表現に改良していくことで十分に改善できると思われる。

「今後も使ってみよう」の評価平均値は3.5であり、「使いやすい」の評価に対し両側検定、有意水準5%で高い評価が得られた。そして「作ってみたい」の評価平均値は3.9であり、こちらは「使いやすい」の評価に対し両側検定、有意水準1%で高い評価を得た。この「作ってみたい」の評価は男子学生(13名)に限定して評価の平均値を見ると4.3となり、男子学生からの評価が特に高かった。今回の研究でシステム製作が可能であり、製作プロセスもわかったので、このようなシステムを製作するような選択の科目を企画しても良いと思われる。このような授業は学生のモチベーションも高く、本学の教育の充実を図ることが期待できるだろう。

以上の評価の結果から、今回開発した対話システムにおける感情表現はユーザーに楽しさを提供できることがわかった。身体的コミュニケーションを用いた対話システムは十分に有用性があると思われる。今回の評価結果から身体的コミュニケーションが効果的だと考える理由は次の通りである。

- 「見やすい」や「使いやすい」の評価が低い場合、一般にシステムの総合評価も低くなりやすいが「面白い」の評価が非常に高く、標準偏差も小さい。
- 「CG」の感情表現が「会話文」の感情表現より有意に高く、CGなどの身体的コミュニケーションを付加することに効果があると考えられる。

実際にシステムを製作するときは、機能をシンプルにして、感情表現もよりわかりやすく改善すると、より良い対話システムになるだろう。

5. システムの応用について

今回の研究成果は、さまざまなシステムと組み合わせることで対話ロボットやCG表示の対話ソフトなどの実現が可能である。今回の結果をもとに、次のような応用例を考えてみた。

- 1人暮らし、核家族向けの対話システムの可能性
アザラシロボット「パロ」(つくば産業技術総合研究所)は視覚、聴覚、触覚、運動感覚を持ち、学習機能により飼い主、環境に合わせて成長するという癒し型ロボットであ

り、様々な場面でセラピー効果もあると報告されている [4]。このようなロボットにユーザーの語り掛けに対して身体的コミュニケーションで応答する対話ソフトを組み入れることで更なる機能の充実が図れるだろう。こういったロボットは必ずしもセラピーを目的としなくても、1人暮らしや核家族向けのペット型ロボットとしても期待できる (図 9)。



図 9 身体的コミュニケーションを備えたペット型ロボットの例

○ 認知症の高齢者の話し相手としての可能性

認知症の高齢者の介護現場では、繰り返される同じ話の相手をする苦勞から対話システムへのニーズが高いと言われている。こうした現場に対話システムやロボットを用いることへの否定論も存在するが [5]、近年、社会問題となっている少子高齢化により今後介護者の人手不足も考えられる。この問題に対しこのようなシステムは、音声認識と合成音声出力を用いることで嫌がることなく認知症の高齢者の話の相手になることができるだろう。さらに、会話のログを記録することが可能なので介護方針の検討にも役立つ。全ての問題の解決にはならずとも、施設で働く人々の労力の軽減に結びつくものだと期待できる。もちろん、前述のペット型ロボットに音声認識や音声出力をつけて、1人暮らし、核家族向けの対話システムとしての利用も考えられる。

○ 聴覚障害者向けの対話システムとしての可能性

今回開発したシステムはキーワード対応のシステムなので、キーワードとなる単語の認識ができれば応答が可能である。ユーザーの手話の単語を画像認識して読み取り、それに対し応答文を示すといった聴覚障害者が手話を使いながらおしゃべりができるという対話システムが考えられる。手話単語の認識はまだ不十分な状況であるが、認識できる単語だけに限定してもシステムの構築が可能であり、前述のペット型ロボットへの応用も期待できる。

○ PC グッズ・コンテンツとしての可能性

PC で文章を書くツール (ワープロ、メモ帳、メールなど) を利用中、入力された文章や受信された文章などに応じて CG など感情表現を加えながら反応してくれる USB グッズなどの開発も図れる (図 10)。



図 10 対話システムを取り入れた USB グッズの例

6. まとめ

今回開発したシステムのように、単なる言葉だけでなく、身体的コミュニケーションを取り入れた対話システムはこれまでに提案した例の他にもさまざまな応用の可能性が見出せる。今後、本研究の成果がより広く、多くの人々に活用され、心豊かな充実した社会の構築に結びついていくことを心から願い、期待している。

参考文献

- [1] 大人の科学マガジン, 大人の科学マガジン Vol.29 AKARI 折紙, 学習研究社, 2010.
- [2] 秋山智俊, 恋するプログラムー Ruby で作る人工無脳ー, 毎日コミュニケーションズ, 2005.
- [3] 奈良先端科学技術大学院大学 松本研究室, "形態素解析システム「茶筌」", <http://chasen.naist.jp/hiki/ChaSen/>, 最新更新 Mar.2007
- [4] 瀬名秀明, 14歳の世渡り術ーロボットとの付き合い方, おしえます. ー, 河出書房新社, 2010
- [5] 広瀬茂男, ロボット創造学入門, 岩波ジュニア新書, 2011.

Development of an Embodied Communication System: “OSHABERI 6256”

NAKAYAMA Mutsumi and NAITO Ichiro

Department of Industrial Information, Faculty of Industrial Technology,
National University Corporation Tsukuba University of Technology

Abstract: Non-verbal forms of communication such as facial expressions and nods play an important role in communication. For this study, we designed a dialogue system that could respond appropriately to an input sentence. We then used this prototype system to examine the effectiveness of these types of embodied communication. The proposed system also responds to physical behavior with CG animation, robots, and LED lights in response to an input sentence. We implemented a prototype system and conducted an experiment to evaluate the effectiveness of these types of embodied communication when we added them to the system. This paper reports the results of our evaluation of the prototype system, and provides a discussion of potential uses for this system based on our results.

Keywords: Embodied communication, Emotion model, Dialogue system