

漢詩添削ソフトによる漢詩知識の獲得方法について

On the knowledge obtain method of Chinese poem by its correction software

石田 勝則*¹
Katsunori Ishida

角 康夫*¹
Sumi Yasuo

西田 豊明*¹
Nishida Toyoaki

*¹ 京都大学大学院情報学研究科
Graduate School of Informatics, Kyoto University

This document describes a discussion on the correction method of the Chinese poem and also gives the scheme to obtain the knowledge of it by the correction support services for the newcomer.

1. はじめに

科学技術論文が論理的表現を重視した表現であるとするれば、漢詩に代表される詩詞は、人間の感情の動きや思想という情念を表現するものであり、その対極にある表現形式といえる。過去に作詩された漢詩の数は100万首を超えといわれ、現在もなお延々と作られ続けている。漢詩は、3000年の歴史を通じて、近体詩または古詩とよばれる表現形式として完成され、今後も発展し続けていくものと思われる。漢詩情報処理を円滑におこなうためには、漢詩の収集、保存、解析、分析、評価を円滑におこなうことができる、漢詩情報処理支援環境が整備される必要がある。本論文は一般の漢詩作詩を志す初心者を対象とした漢詩添削ソフトウェアをウェブ上に公開し、初心者に対する漢詩添削活動を通じて、漢詩に関する知識を蓄積する仕組みについて論じる。

2. 漢詩知識獲得の基本フレーム

よりよい漢詩を読むためには、立派な漢詩を鑑賞することが重要だと言われる。名作といわれる漢詩は、長い淘汰の歴史を経て、人々によって伝承されてきたものであり、今尚鑑賞に堪えるものである。これらの作品を読みかつ鑑賞することは、新たに漢詩を作る初心者にとって、有効な学習方法であると言われている。初心者といえども、作られた漢詩は、それが他人によって読まれ、且つ鑑賞されて始めて意味をもつ。この漢詩添削システムは、初心者が作詩した漢詩に対し、文法上の添削サービスを提供し、一次評価を行うとともに、その作品をウェブ上に公開し、多くの鑑賞者が評価した結果を整理・蓄積し、漢詩知識を蓄積する仕組みを提供するものである。

3. 漢詩添削処理について

本漢詩添削ソフトウェアの漢詩規約は、参考文献に示す[太刀掛 90] 太刀掛重雄: だれにでもできる漢詩のつくり方, 呂山詩書刊行会, (1990) に従った。以下に使用した漢詩規則を示す。

3.1 漢詩添削ソフトウェアの漢詩規則

(1) 句の構成

句は五言絶句では二字・三字の五文字、七言絶句では二字・二字・三字の七文字で構成され、一つの絶句は一句(起句)、二句(承句)三句(転句)、四句(結句)の四句で構成される。

(2) 平仄チェック

すべての漢字は四声により平字・仄字のいずれかに属する。絶句には平字の二字で始まる平起式と、仄字の二字で始まる仄起式がある。

七言絶句(仄起式)の平仄配列規則

一句
二句
三句
四句

ここで、は平字、は仄字をあらわす。

(3) 平仄公式の例外規則

- 一三五不論(1, 3, 5字目は平仄どちらでもよい)
- 二四不同二六対の原則
- 孤平(五言の二字目と七言の四字目の平字が仄字で挟ること)の禁止
- 挟み平の容認(平起式の転句の下三字の を に変えてもよい)
- 下三字が連続して平字または仄字になることの禁止
- 拗体(平起式・仄起式それぞれにおいて一・二句と三・四句の平仄を入れ替えてよい)の容認

(4) 押韻チェック

五言絶句、七言絶句の句形式に応じた音韻規則のチェック
五言絶句の場合は二・四句、七言絶句の場合は一・二・四句の最後の文字が韻(音韻が同じ響きをもった漢字の集合であり、106種類ある)を踏んでいること

(5) 詩語チェック

漢詩は中国の詩であるので、和製の漢字語句は和臭とよばれ使用できないことになっている。漢和辞典にある漢語かどうかのチェックが必要となる。

二文字漢字、三文字漢字はそれぞれ同じ文字数の詩語辞典と比較することによりチェックする。

以上が今回採用した漢詩規則であり、漢詩添削ソフトウェアの核文法に相当するものである。尚、例外規則については添削を通じて事例を収集し、追加適用を検討する。

3.2 漢詩添削処理の事例考察

今回実装した漢詩添削システムの機能は統語論レベルの初歩的段階のシステムであり、より高度の機能を備えるために、複数の経験者(人間)による漢詩の添削事例について考察する。添削例:

形式 七言絶句(仄起式)
韻 (平)庚
詩題 大衆禅堂

<原文 / 平仄規則>	<読み下し文>
夕刻 ^{*4} 禅 ^{*1} 堂 ^{*3} 靈 ^{*3} 氣 ^{*2} 増 ^{*2}	夕刻草堂靈氣増し
経文 ^{*4} 齊 ^{*4} 唱 ^{*4} 熱 ^{*4} 心 ^{*4} 堂 ^{*4}	経文齊唱熱心に営む
精神 ^{*1} 統一 ^{*3} 座 ^{*3} 禅 ^{*3} の 業 ^{*3}	精神統一座禅の業
不 ^{*1} 惑 ^{*1} 思 ^{*1} 留 ^{*1} 妙 ^{*1} 麗 ^{*1} 情 ^{*1}	惑わず思いを留める妙麗の情
<添削後の文>	<読み下し文>
夕暮 ^{*1} 禅 ^{*1} 堂 ^{*1} 靈 ^{*1} 氣 ^{*1} 成 ^{*1}	夕暮禅堂靈氣成し
経文 ^{*1} 朗 ^{*1} 誦 ^{*1} 洗 ^{*1} 心 ^{*1} 清 ^{*1}	経文朗誦して洗心清し
精神 ^{*1} 統一 ^{*1} 只 ^{*1} 趺 ^{*1} 座 ^{*1}	精神を統一し只趺座すれば
煩 ^{*1} 悩 ^{*1} 超 ^{*1} 然 ^{*1} 妙 ^{*1} 麗 ^{*1} 情 ^{*1}	煩悩に超然たり、妙麗の情

この漢詩の原文は、まだ漢詩をはじめて間もない漢詩愛好家が早朝、ある禅寺において公開されている座禅業に参加して読んだものの投稿である。

- まず禅^{*1}の漢字が重複して使用されている。(重複使用の禁止)
- 起句の押韻部の増^{*2}が(平)庚でない。(押韻規則違反)
- 仄起式の七言絶句の平仄順序は一句と三句の^{*3} ^{*3}が平仄規則から外れるが、付加規則の2)一三五不問ルールにより問題がないことがわかる。(以上は漢字字典に示される漢字の平仄・音韻データベースを準備することにより、検出が可能である。)
- “夕刻^{*4}”、“熱心^{*4}”は和泉と呼ばれ、漢語でない。和製漢字は原則として漢詩に使用できないので、平仄に注意しながら、同意の漢語に置き換えなくてはならない。そのためには和製漢語の収集と、同義または類義漢語の対応データベースが必要である。
- 核文法を用いた統語論によるチェックを行う場合に厄介なのは、漢字の中にも同一文字で平・仄両用のものがあることである。(その場合意味も異なる場合が多い。) また多品詞多義語が含まれることにより、統語解析上の曖昧さを生ずることになるが、自然言語処理と同様、意味解析や統計的処理で絞りこみ、そのうえでさらに平仄チェックをかける必要がある。
- 次に意味論上の誤りについて考察する。二句目の“経文齊唱”は統語解析では正しい表現であるが、経文は“朗誦する”ものであり、“齊唱する”は適当ではない。これは漢語の用法上の不都合である。この誤りを指摘するためには、漢語の照応関係を意味解析レベルで関連づけた知識ベースが必要である。
- 次に作者が漢詩で表現したかった主題(意図)について考える。この漢詩の意境(作者の真意)は“座禅をすることによって、妙麗の美人の姿にも気をそらさず、煩悩を払いのけるほど精神統一ができた。”ということであるから“妙麗情”は動かせない。絶句では一般に結句(4句目)のどこ

ろに作者の意図が隠されているといわれている。この点に留意して(作者の意図を大切に)類似漢語で置き換える必要がある。平仄に注意して“座禅業”は“只趺座”に置き換わった。

- 次に詩文の意味に矛盾はないかを考える。作者の結句の詩文では“惑わず思いを留める妙麗の情”となっており、惑わず美人の姿を心に留めることになり、作者の意境の逆の意味になってしまう。経験ある鑑賞者は作者の意図を正確に表現していないことを即座に指摘できる。

添削ソフトウェアは、はたしてこれらの誤りを的確に指摘できるであろうか? そのためには添削ソフトウェアは“座禅は煩悩を払うために行うものである”という常識知識や、“読経、や精神統一”の意味を理解した上で、結句の“不惑思留妙麗情”を解析し誤りを指摘することが求められる。そのためには言語学の成果である言外の意味を推論する語用論[内海 03]や前段の句で読まれた文の意味を踏まえた意味解析をおこなう、関連性理論を用いた認知語用論[松井 03]の活用が考えられる。形態論、統語論、意味論、語用論を統合した漢詩文全体の意味を総合的に捉えるメカニズムを備えた、より高度な漢詩添削システムの実現が課題となる。

4. 漢詩情報処理の目標

添削システムに誤りを指摘するだけでなく、原文を修正する機能までを含めれば、漢詩添削システムは、漢詩作詩支援システムに発展することは明らかである。これには意味解析から出発し、詩語の知識ベースを活用したヒューリスティックなアプローチが考えられる。[西田 88]が指摘するように、言語解析やシステムの枠組みができて、適切な文法や辞書がなければ魂のない抜け殻のようなものになりにくい。過去の詩人の漢詩作品の鑑賞が漢詩作詩のよい手本となると先に指摘したが、漢詩データの収集、分類、整理、および辞書データベースへの反映が効果的に進められる仕組みを提供することが前提となる。漢詩に使われている漢語用法の収集と解析は、統合的なパーシングを行う場合の優先意味論的知識を蓄積する立場からも有効なアプローチである。以上の分析から、漢詩情報処理の目標として次の三つのシステムを設定した。

漢詩評価システム

文法上の不都合や、親和性のない表現、語法の誤りなどのマイナス評価と、漢詩の形式に合致した点や情感の表現方法として優れた点などのプラス評価を採点し、完成した漢詩を多角的に評価するソフトウェア

漢詩作詩支援システム

漢詩のモチーフを与えれば、モチーフに沿って詩語を選択し、作者とインタラクティブに会話しながら、漢詩を作詩することを可能にするソフトウェア

漢詩読み上げシステム

今後の課題はユーザの気持ちを汲み取って気の利いた、暖かみのある応答を返す漢詩添削システムの実現である。

参考文献

- [太刀掛 90] 太刀掛重雄: だれにでもできる漢詩の作り方, 呂山詩書刊行会, (1990)
- [西田 88] 西田 豊明: 自然言語処理入門, オーム社, (1988)
- [内海 03] 内海 彰: 言外の意味のコミュニケーション - 語用論概説 - , 人工知能学会誌, Vol.18 No.3, pp.337-345, (2003)
- [松井 03] 松井 智子: 関連性理論 - 認知語用論 - , 人工知能学会誌, Vol.18 No.5, pp.592-602, (2003)