

Automating Idiom Translation with Cross-Lingual Natural Language Generation Grounded In Semantic Analyses Using Large Language Models

Ming Qian

Natural Language Processing Group
Human-Centered AI, Charles River Analytics

mqian@cra.com
qianmi@gmail.com

Challenges of Idiom Translation

Idioms are multi-word expressions whose meaning cannot be directly derived from the meanings of their component words

- "kick the bucket" means "to die"

Translating idioms from one language to another requires learning the mapping between the idiom in the source language and its corresponding idiomatic expression in the target language

- Literal word-for-word translation is not good
- Acquired and saved as translation lexicon

Challenges:

- Idioms exhibit varying degrees of semantic transparency
- Cross-language differences in idiom usage and connotations add complexity to the translation task

Can we leverage large language models to mining possible solutions following a "standard" step-by-step procedure?

Adding the Semantic Analyses Step

1. Semantic analysis using ontology [1] or FrameNet [2] on a Chinese idiom to identify key concepts and relationships such as action, purpose, outcome, and context.
2. Generation of multi-word English expressions that reflect similar, narrowed-down, or broadened concepts and relationships.
3. Selection of the top candidate for an English idiom that closely matches the meaning of the Chinese idiom.

Reference:

[1] Natural Language Ontology, Stanford Encyclopedia of Philosophy, <https://plato.stanford.edu/entries/natural-language-ontology/>

[2] Berkeley FrameNet Website, <https://framenet.icsi.berkeley.edu/>

[3] Strategies for Translating Chengyu (Chinese Idioms), Jim Jones, <https://www.ata-divisions.org/CLD/strategies-for-translating-chengyu/>

Chinese idiom "破釜沉舟" direct output from GPT-40

Translate the Chinese idiom "破釜沉舟" into English

The Chinese idiom "破釜沉舟" (pò fǔ chén zhōu) translates to "break the cauldrons and sink the boats" in English. It means to commit oneself to a course of action by removing any possibility of retreat, often used to describe someone making a determined and irreversible decision. The phrase originates from a historical event where a general ordered his troops to destroy their cooking pots and boats to force them to fight to the end, knowing there was no way back.

Translate the Chinese idiom "破釜沉舟" into a corresponding English idiom

The corresponding English idiom for "破釜沉舟" (pò fǔ chén zhōu) would be "burning one's bridges." Both idioms convey the idea of making a decision that eliminates the possibility of retreat, thereby committing fully to the chosen course of action.

Chinese idiom "破釜沉舟"

Ontology Analysis

Ontology analysis involves looking at the concept and its relationships with other concepts. The literal translation of "破釜沉舟" is "break the pots and sink the boats." This idiom originates from a historical event during the Qin Dynasty when a general named Xiang Yu ordered his troops to destroy their cooking pots and boats after crossing a river into enemy territory. This was done to eliminate any possibility of retreat, thereby motivating his troops to fight forward against the enemy as their only option for survival.

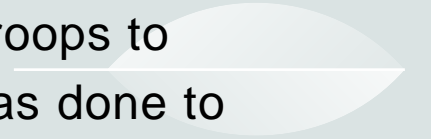
Key Concepts and Relationships:

Action: Destruction (of the boats and pots)

Purpose: Eliminate means of retreat

Result: Commitment to a single course of action, i.e., to fight forward

Context: Military strategy, determination, and resolve



Chinese idiom "破釜沉舟"

FrameNet is a project that catalogs how words and phrases fit into different frames (or scenarios) and describes the roles and relationships within these frames.

For "破釜沉舟," the relevant frames might include:

Destruction: This frame involves an agent (the subject carrying out the destruction) and the items being destroyed (pots and boats in this case). It involves intentional action aimed at rendering something unusable or ineffective.

Commitment: This frame would involve committing to a plan or action with determination, where the commitment is demonstrated through some significant act (in this case, the destruction of retreat options).

Resolve: A subframe of Commitment, which deals specifically with a firm decision to act in a particular way, often in the face of obstacles.

Elements of FrameNet Analysis:

Agent: General Xiang Yu and his troops

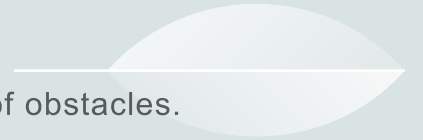
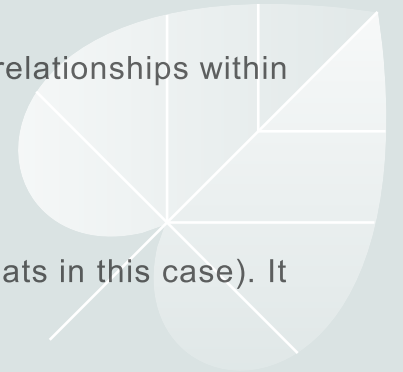
Theme: The pots and boats

Purpose: To ensure no possibility of retreat

Means: Destruction

Setting: Military encampment during a military campaign

Through these analyses, "破釜沉舟" conveys a broader, metaphorical meaning that extends beyond its literal sense. It symbolizes making a decisive and irreversible decision with a clear focus on achieving a goal or objective, often used in contexts where abandoning all alternative options or potential escape routes underscores the seriousness and determination of the commitment.



Chinese idiom "破釜沉舟"

Here are the top 5 candidates for an English idiom that closely match the meaning of the Chinese idiom "破釜沉舟" based on the frame and ontology analyses:

- **Burn one's boats**
- **Burn one's bridges**
- **Cross the Rubicon**
- **No turning back**
- Go all in

Human Expert's Translation

Reference: Strategies for Translating Chengyu (Chinese Idioms), Jim Jones, <https://www.ata-divisions.org/CLD/strategies-for-translating-chengyu/>

- **Burning one's boats**
- Burning the bridges
- Point of no return
- crossing the Rubicon



Chinese idiom "刀山火海" direct output from GPT-40

Translate the Chinese idiom "刀山火海" into English

The Chinese idiom "刀山火海" can be translated into English as "a mountain of swords and a sea of flames." This phrase metaphorically describes a situation filled with extreme danger or severe challenges, indicating that someone is willing to go through the most dangerous and difficult situations to achieve a goal or demonstrate loyalty.

Translate the Chinese idiom ""刀山火海"" into a corresponding English idiom

The Chinese idiom "刀山火海" (dāo shān huǒ hǎi) literally translates to "mountain of knives and sea of fire." It metaphorically describes a situation that is extremely dangerous, difficult, or filled with challenges. A corresponding English idiom would be "going through hell and high water" or "facing insurmountable obstacles."

Chinese idiom "刀山火海"

Ontology Analysis

Literal Translation

"刀山火海" translates literally to "mountain of knives and sea of fire."

Ontology analysis involves breaking down the components of the idiom and understanding its semantic relationships. Here's a breakdown:

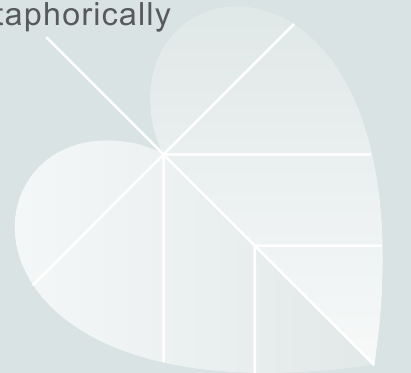
刀 (Dāo) - Knife: Represents danger, sharpness, and potential harm.

山 (Shān) - Mountain: Implies something large, difficult to overcome, or insurmountable.

火 (Huǒ) - Fire: Symbolizes danger, intensity, and destruction.

海 (Hǎi) - Sea: Represents something vast, difficult to navigate, and possibly overwhelming.

These elements combine to create a vivid metaphor for a situation that is both perilous and challenging, where the person is metaphorically climbing a mountain of knives and crossing a sea of fire. This idiom conveys the idea of extreme hardship and adversity.



Chinese idiom "刀山火海"

FrameNet is a theory that focuses on the semantic frames, or conceptual structures, that words and phrases evoke. Let's apply this to "刀山火海":

Frame: Danger

Core Elements:

Dangerous Object (Knives/Fire): The "刀" (knife) and "火" (fire) are central elements that evoke the concept of danger.

Environment (Mountain/Sea): The "山" (mountain) and "海" (sea) provide a setting that amplifies the sense of peril.

Peripheral Elements:

Experiencer: The implied person or entity facing the danger.

Outcome: The potential for harm or disaster if the challenge is not overcome.

Frame: Challenge/Adversity

Core Elements:

Obstacle (Mountain/Sea): The mountain of knives and sea of fire represent significant obstacles or challenges.

Effort: The need for extraordinary effort to overcome these challenges.

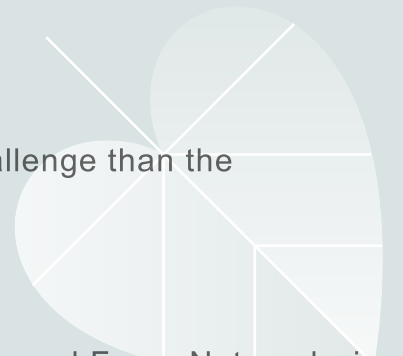
Peripheral Elements:

Motivation: The reason why the person is facing such dangers, which could be inferred from context.

Resolution: The possible overcoming of the challenges, though the idiom focuses more on the existence of the challenge than the resolution.

Summary

The idiom "刀山火海" evokes a frame of extreme danger and challenge, combining elements of peril and adversity. The ontology and FrameNet analysis reveal that the idiom encapsulates the essence of facing overwhelming difficulties, where success is uncertain, but the stakes are high.



Chinese idiom "刀山火海"

Here are the top 5 candidates for an English idiom that closely match the meaning of the Chinese idiom "刀山火海" based on the frame and ontology analyses:

1. **Run the gauntlet**
 - This idiom means to go through a series of difficulties or dangers to reach a goal, similar to facing a "mountain of knives and sea of fire."
2. **Walk through fire**
 - This idiom conveys the idea of enduring extreme challenges and difficulties, akin to navigating through a perilous situation.
3. **Face the music**
 - While this idiom primarily means to confront the consequences of one's actions, it can also imply facing difficult situations head-on with courage.
4. **Go through hell and high water**
 - This idiom means to endure severe trials and tribulations, reflecting the idea of facing extreme danger and difficulty.
5. **Into the lion's den**
 - This idiom means to enter a very dangerous or hostile situation, similar to confronting a mountain of knives and a sea of fire.

Human Expert's Translation

Reference: Strategies for Translating Chengyu (Chinese Idioms), Jim Jones, <https://www.ata-divisions.org/CLD/strategies-for-translating-chengyu/>

- **Literally meaning "mountain of knives and sea of fire,"**



Chinese idiom “抛砖引玉” direct output from GPT-40

Translate the Chinese idiom "抛砖引玉" into English

The Chinese idiom "抛砖引玉" literally translates to "throwing a brick to attract jade." It means to offer something of lesser value in the hope of receiving something more valuable in return. It's often used metaphorically to describe someone who presents a simple or modest idea or effort to inspire or prompt others to contribute something better or more refined.

Translate the Chinese idiom ""抛砖引玉"" into a corresponding English idiom

The Chinese idiom "抛砖引玉" can be translated into the English idiom "to break the ice." Both idioms convey the idea of making the first move, often with something modest or simple, to encourage others to contribute something more significant or valuable.

Chinese idiom "抛砖引玉"

Ontology Analysis of the Chinese Idiom "抛砖引玉"

1. Basic Components:

抛 (pāo): To throw or to cast.

砖 (zhuān): Brick, a symbol of something of lesser value or commonness.

引 (yǐn): To attract or lead.

玉 (yù): Jade, a symbol of something valuable, precious, or of higher quality.

2. Ontological Categories:

Entities:

Brick (砖): Represents something of lesser value, often used metaphorically to denote a modest or insignificant offering.

Jade (玉): Represents something of higher value, often used metaphorically to denote a valuable response or outcome.

Actions:

Throwing (抛): An action initiating a process or interaction.

Attracting (引): An action that results from the initial action, leading to a response.

3. Relationships:

Causal Relationship: The act of throwing a brick (a simple or humble offering) leads to attracting jade (a valuable or significant return).

Hierarchical Relationship: The brick is ontologically inferior to the jade, which represents the valuable outcome that one hopes to achieve.

4. Metaphorical Mapping:

Source Domain: Throwing a brick (making a modest effort or contribution).

Target Domain: Attracting jade (inspiring or eliciting a valuable response or contribution from others).



Chinese idiom "抛砖引玉"

Frame Analysis Based on FrameNet Theory

Frames Involved:

Giving Frame: The "throwing" of a brick implies an initial act of offering or giving, even though the value of the offering is modest.

Receiving Frame: The "attracting" of jade implies the reception of something valuable, indicating a positive outcome resulting from the initial act.

Value Judgment Frame: The contrast between the brick (less valuable) and the jade (more valuable) involves a judgment of value, highlighting the expectation of a more significant return for a smaller investment.

Frame Elements:

Agent: The person or entity that throws the brick, symbolizing the initiator of the action.

Object: The brick, symbolizing the initial offering or contribution.

Recipient: The person or entity who is expected to respond with the jade, symbolizing the valuable contribution or response.

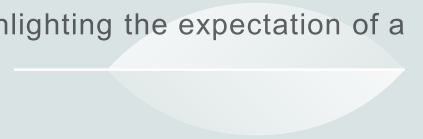
Result: The acquisition of jade, symbolizing the valuable outcome or reward resulting from the initial action.

Frame Relations:

Inheritance: The "Giving" frame inherits from a broader frame of "Transaction" where an exchange is anticipated.

Subframe: The "Receiving" frame can be seen as a subframe of the broader "Transaction" frame, specifically focusing on the outcome or response.

Perspective: The idiom takes the perspective of the initiator (the one throwing the brick) and anticipates a response from others (attracting jade).



Chinese idiom "抛砖引玉"

Here are the top 5 candidates for an English idiom that closely match the meaning of the Chinese idiom “抛砖引玉” based on the frame and ontology analyses:

"To cast a small stone to attract a gem"

"To plant a seed and harvest a tree"

"To give a penny and get a pound"

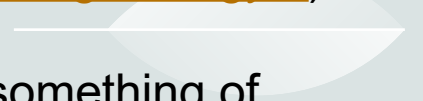
"To break the ice for bigger ideas"

"To offer a spark to ignite a flame"

Human Expert's Translation

Reference: Strategies for Translating Chengyu (Chinese Idioms), Jim Jones, <https://www.ata-divisions.org/CLD/strategies-for-translating-chengyu/>

"to get the ball rolling by starting with something of average value."



Chinese idiom “不明觉厉” direct output from GPT-40

Translate the Chinese idiom “ ” into English

The Chinese idiom "不明觉厉" (bù míng jué lì) can be translated into English as "It's unclear but impressive" or "I don't understand it, but it seems awesome." This idiom is often used to describe something that seems profound, complex, or advanced, but the speaker doesn't fully understand it. Despite the lack of understanding, the speaker can sense its significance or feels that it must be impressive.

Translate the Chinese idiom “不明觉厉” into a corresponding English idiom

A corresponding English idiom for “不明觉厉” (bù míng jué lì) could be "It's all Greek to me, but it sounds impressive." This idiom captures the idea of something being difficult to understand (like a foreign language) but still appearing significant or impressive.

Chinese idiom “不明觉厉”

FrameNet is a linguistic theory that examines how words evoke certain types of knowledge structures, known as "frames." Frames are conceptual structures that describe various situations, events, or objects and the participants involved.

For the idiom "不明觉厉" (bù míng jué lì), the analysis can be broken down as follows:

1. Core Frame: Lack of Understanding

1. Participants:

1. **Cognizer:** The individual who is experiencing the lack of understanding.
2. **Phenomenon:** The subject or content that is not understood.

2. Frame Elements:

1. **Lack of Clarity:** The "不明" part of the idiom indicates that the subject is not clearly understood by the cognizer.

2. Core Frame: Perception and Evaluation

1. Participants:

1. **Evaluator:** The individual who perceives and evaluates the subject.
2. **Phenomenon:** The subject or content being perceived and evaluated.

2. Frame Elements:

1. **Subjective Perception:** The "觉" part of the idiom focuses on the personal feeling or reaction to the unclear subject.
2. **Positive Evaluation:** The "厉" part of the idiom provides an evaluative judgment that the unclear subject is impressive or awesome.

3. Integration of Frames:

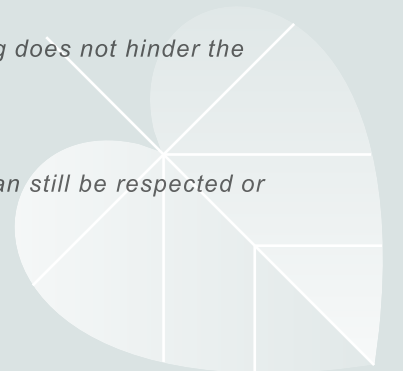
1. *The frames of **Lack of Understanding** and **Perception and Evaluation** are integrated to form the overall meaning of the idiom. The lack of understanding does not hinder the positive evaluation; rather, the positive evaluation emerges precisely because the subject is perceived as complex or profound.*

4. Cultural Frame:

1. *In a broader cultural context, the idiom reflects a value placed on expertise, complexity, and profundity. Even when something is not fully understood, it can still be respected or admired for its apparent depth or significance. This might suggest an underlying cultural appreciation for mystery or complexity.*

5. Frame Relations:

1. **Causality:** *The lack of understanding causes a subjective perception.*
2. **Connotation:** *The subjective perception leads to a positive connotation (impressive or awesome).*



Chinese idiom "不明觉厉"

Here are the top 5 candidates for an English idiom that closely match the meaning of the Chinese idiom “不明觉厉” based on the frame and ontology analyses:

"I don't get it, but it's impressive."

"It's over my head, but it's awesome."

"Sounds complicated, but it must be brilliant."

"Looks mysterious, but it's clearly genius."

"I'm lost, but it's obviously genius."



Chinese idiom "男女有别"

Here are the top 5 candidates for an English idiom that closely match the meaning of the Chinese idiom "男女有别" based on the frame and ontology analyses:

"Men are from Mars, women are from Venus."

This idiom emphasizes the inherent differences between men and women, suggesting that they are almost from different worlds in terms of their characteristics, behavior, and thinking.

To each his own.

"While this idiom is more general, it can be applied to suggest that different genders have their own roles, preferences, or behaviors, which should be respected.

"Vive la différence!"

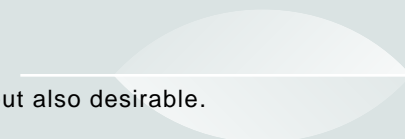
This French phrase, often used in English, celebrates the differences between the sexes, suggesting that these differences are not only natural but also desirable.

"A woman's place is in the home."

Although outdated and controversial, this idiom reflects traditional gender roles, implying that men and women have distinct societal roles.

"Separate spheres."

This phrase historically refers to the idea that men and women occupy different roles or spheres in life—men in public, professional life, and women in private, domestic life.



Conclusion

- The performance was on par with human experts
- By analyzing the semantics, a core aspect of human intelligence, the concepts, properties, and logical rules informing the system's decisions can be understood to enhance translation
- Additionally, semantic reasoning can simulate the LLMs' decision process transparently by applying the same logical inferences over the semantic framework generated by LLMs





Thank you

Ming Qian

919-523-1320

mqian@cra.com

qianmi@gmail.com