

Generating Instrumental Expressions in a Multilingual Question-Answering System



Monojit Choudhury (IIT Kharagpur, India)

Elixabete Murguia (Univ. Deusto, Bilbao, Spain)

Sudeshna Sarkar (IIT Kharagpur, India)

Veronique Moriceau (IRIT, France)

Asanee Kawtrakul (Kasetsart univ., Thailand)

Patrick Saint-Dizier (IRIT, France).

Instrumental Questions

- Factoid questions that ask about the instrument that is/has been used to accomplish a particular task
- Examples:
 - What is used for cutting glass?
 - What do you open a lock with?
 - Which algorithm can we use to sort an array of numbers?
 - How do you travel from Paris to London?

Motivation

- Instrumental questions are most common type after the who/where/when questions over web
- Very important in the domains of
 - Technology
 - Travel and tourism
 - Cooking etc.
- Demands a special treatment

Objectives

- Analysis of the notion of **instrumentality** and its surface realization in different languages
- A classification system for Instrumental questions
- Decision graphs for 5 languages
- question analysis and answer generation in a Multilingual QA System

What is Instrumentality ?

‘an instrument is an artifact, or a set of artifacts, that are instrumental in accomplishing some end’ (WordNet)

- non-volitional
- controlled by an agent
- roles of agent and instrument w.r.t. action is quite diverse.

Languages

- Bengali
- French
- Hindi
- Spanish
- Thai

- English (no decision graph)

Difficulties

- almost everything can be an instrument for some action (e.g. cut glass with a hammer)
- considerable overlap with other notions: causality, manner, paths, etc.
- considerable variations in surface forms depending on languages (prepositions, postpositions, affixes, etc.)
- large number of metonymies and metaphors (write with your heart).

Outline of the QA system

1. The query presented in Hindi (or any other Languages)
3. Query is processed to identify the
 1. **Task** for which the instrument is queried for
 2. Identification of query **type**
4. Search for the (sentences with) **Task** in documents in many languages

Outline of the QA system

1. Analyze the search results for responses
3. Cluster the responses and categorize according to the prototypicality of the instruments
5. Combine responses to generate the answer
7. Present the answer in Hindi using appropriate NLG

Conceptual Parameters

- Concrete vs Abstract Instruments
 - B: *diye* vs. *dwara*
 - T: *duai* vs *tam*

- Subclasses of Concrete instruments:
 - instrument is a recipient (S: *en*, B: *-e kare*) to carry an object along a certain trajectory,
 - instrument is part of the body (T: *kap*),
 - goal difficult to reach (S: a base *de*)

Conceptual Parameters

- **Manner**: some instruments are close to manners and are realized as such,
 - e.g. T: **khian duai muek daeng** (write with red ink)
- **Causal and Agentive**:
 - F: `a cause de,
 - H: **ke zariye** (agentive), vs. **ki vajah se** (causal instrument)

Conceptual Parameters

□ Paths and sources

- B: -e and -te (paths where the agent has no control) vs. diye and dhare (involve control)
- T: source need a special mark, chak

□ Means of transport

- B: kare, -te kare or -ya kare

Conceptual Parameters

- Orientation:
 - indicating a positive (thanks to, T: **khop khun**)
 - negative (S: **por culpa de**)

- Language level: some marks are proper to formal language,
 - H: **se** vs. **dwaaraa**

Generating Instrumental Expressions

- Ontological data resolves
 - Abstract vs. Concrete
 - Means of transport
- Thematic structure and selectional restriction
 - Path
- Global positive/negative orientation
 - Orientation
 - Language level

Role of Corpus

- Manner instruments in French (à)
- Causal and Agentive instruments (are more commonly found to be the subjects)
- Prototypicality of instruments for tasks
- The preferred surface realization in case of multiple possibilities
 - B: **kare** and **-e kare** used for transport but
 - * **jAhAj kare yAo** ✓ **jAhAje kare yAo**
 - ✓ **gAri kare yAo** ✓ **gArite kare yAo**

Realization in Bengali

- Suffixes: -e
- Postpositions: diYe, dwArA, kare
- Combination: -ke diYe, -e kare, -r dwArA
- Special words (usually nouns+inflection, but can be treated as postpositions also): -r sAhAyye, -r janya, -r kR^ipAYe, ...
- Action instrumentalized using the non-finite form of the verb: verb+e (geYe, bale)
- No marker for some verbs (peTa, khelA)

Bengali

Decision Graph

Path:

— Alternative paths exist: *-e, diYe, dhare*

— No alternatives: *-e*

Means of transport: *-e, kare, -e kare, diYe*

Causal:

— Positive orientation: *-r kripAYe, -r kalyANe*

— Neutral orientation: *-r janya, -r phale*

Agentive:

— Specific Agent: *-ke diYe, -r dwArA*

— Non-specific Agent: *diYe*

Indirect instrument: *-r sAhAyye, -r sahayoge*

Resource Instrument: *-e, diYe, khATiYe*

Direct Instrument:

— Verb is “play” or “beat”: null

— Instrument can *contain* the object: *-e, -e kare*

— Agent has control: *diYe*

— Agent has no control: *-e*

Medium: *-r mAdhyame*

Action instrumentalized: Verb-*e* (non-finite)

Hindi

Decision Graph

Path: *se*

Means of transport:

— Animals: *se, pe/para*

— Others: *se, me.n*

Causal:

— Positive orientation: *ke badaulata*

— Neutral orientation: *le lie, ke vAste, ke nAte, ke kAraNa*

Agentive:

— Specific Agent: *ke dwArA, se*

— Non-specific Agent: *se, dwArA*

Indirect instrument: *ke sahAre, ke sAtha, sahita*

Resource Instrument: *dekara, lagAkara, se*

Direct Instrument: *se*

Medium: *ke mAdhyama se, ke zariye*

Action instrumentalized: Verb-*ke/kara* (non-finite)

French

Decision Graph

Path:

- Spatial: *par*
- Temporal: *avec*
- Source: *de*

Means of transport:

- Means, path: *par, en*

Causal: *à cause de*

Manner: *à, en*

Indirect + abstract instrument:

- Agentive: *au moyen de*
- Other cases: *avec*

Direct Instrument:

- Container: *en*
- Focus *à l'aide de*
- Lack of prototypicality: *avec*

Orientation:

- Positive: *grâce à*
- Negative: *à cause de*

Spanish

Decision Graph

Path:

- Spatial: *por, a través de*
- Temporal: *a través de, con*
- Source: *de* + determiner

Means of transport:

- Means, path: *por*
- Container: *en*

Causal: *a causa de*

Manner: *a, en*

Indirect + abstract instrument:

- Agentive: *mediante, por medio de*
- Other cases: *con, por medio de, mediante*

Direct Instrument:

- Container: *en*
- Focus *con la ayuda de*
- Difficulty to realize: *a base de*
- Lack of prototypicality: *por medio de, mediante*

Orientation:

- Positive: *gracias a*
- Negative: *por culpa de*

Thai

Decision Graph

Path:

— Direct: *tam*

— Metaphorical or channel: *thang*

— Source: *chak*

Means of transport: *doi, thang*

Causal: *duai*

Manner: *duai, doi*

Indirect + abstract instrument:

— Agentive: *doi*

— Other cases: *tam*

Direct Instrument:

— Default: *duai*

— Part of the body: *kap*

Positive orientation: *khop khun krap* (Masculine)

/kah (feminine)

Thank You

