

Investigating Conversational Search Behavior For Domain Exploration

Phillip Schneider, Anum Afzal, Juraj Vladika, Daniel Braun and Florian Matthes

{phillip.schneider, anum.afzal, juraj.vladika, matthes}@tum.de, d.braun@utwente.nl

Motivation

- Conversational search is an emerging information retrieval paradigm, replacing classic systems by interactive dialogues with intelligent agents
- This study investigates human conversational search behavior in exploratory information-seeking contexts

Experimental Design

- Laboratory study with 26 students participating in open-ended search conversations for exploring domain-specific tabular datasets
- Participants were assigned information seeker and information provider roles
- Interactions were restricted to a text chat interface

Collected Dialogue Corpus

- To analyze conversational behavior, we collected dialogues for five domains in a laboratory study

Domain	No. of Messages
Geography	169
History	98
Media	156
Nutrition	153
Sports	93

Table 1: Message count for explored domains

Information Seeker

Chat Interface

Information Provider

Domain Database

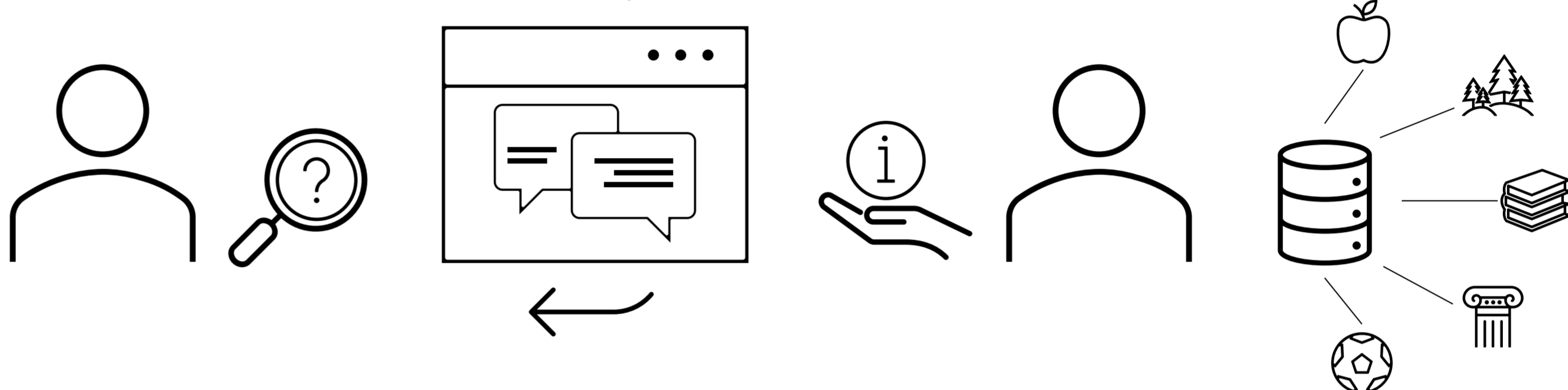


Figure 1: Overview of the experimental design

Analysis Results

- Similar distribution of speech acts across the domains (Figure 2), and mixed-initiative interactions through requesting (directives) and providing (constatives) information items
- Extracted sequence flow of most frequent dialogue acts (Figure 3) reveals three main loops: (1) information requests, (2) item descriptions, as well as (3) aggregating and grounding dialogue acts

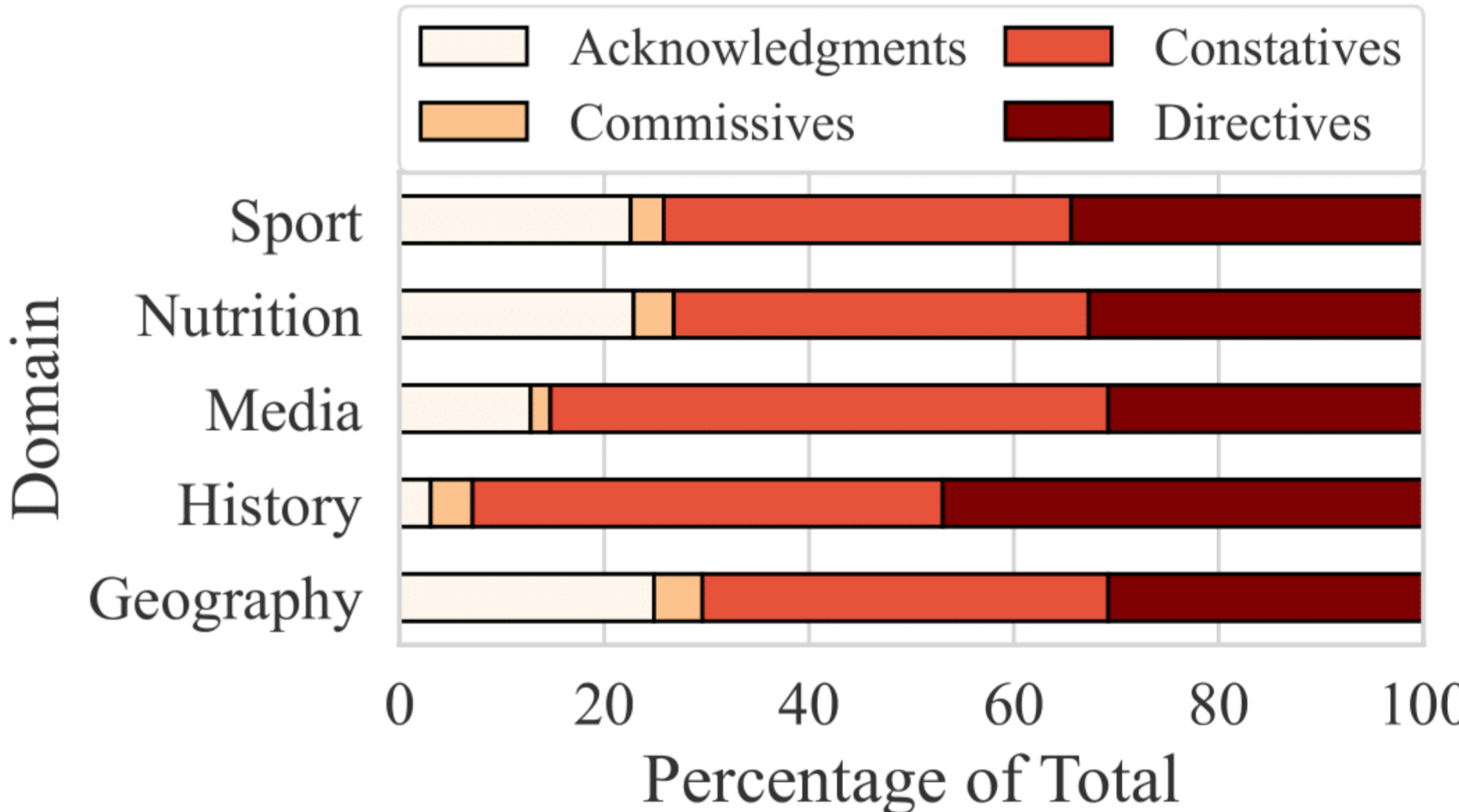


Figure 2: Distribution of speech acts across five domains

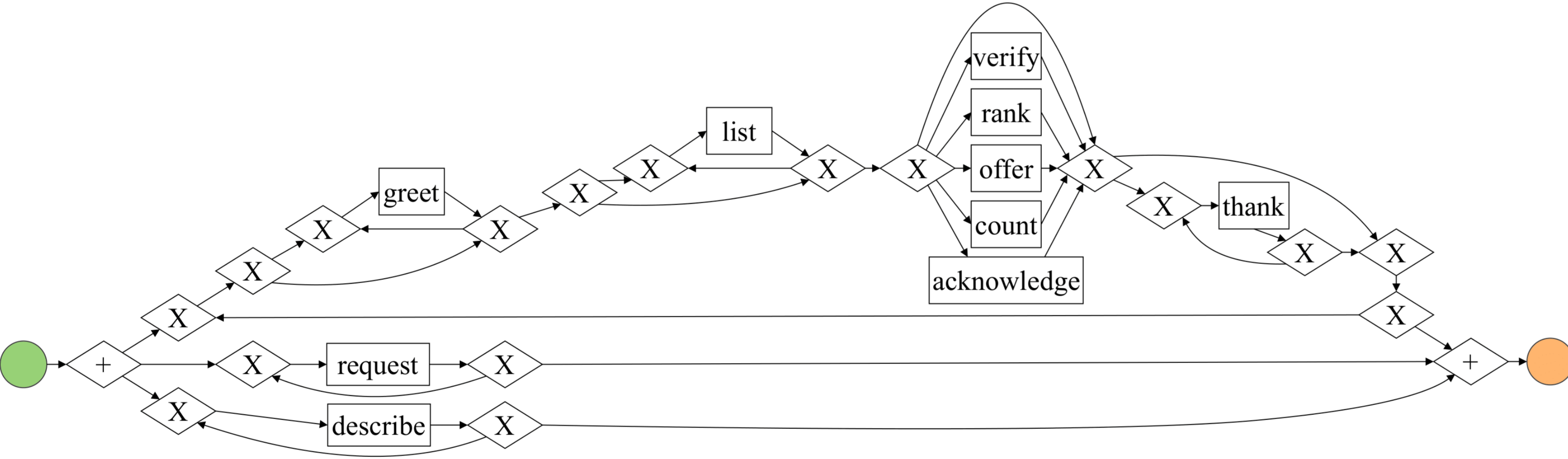



Figure 3: Process model of extracted dialogue sequence flow

- Aggregating dialogue acts, such as “list”, “rank”, or “count”, were crucial to understand the dataset’s dimensionality
- Information seekers preferred: brief dataset summary, detailed specific answers, and question suggestions

Conclusion & Future Work

- Identification of general search-focused dialogue sequence in five domains
- Brief introduction of metainformation at the start of the session is an essential requirement for engaging in deeper exploration
- Dialogues mostly rely on directives and constatives, although information seekers use a variety of dialogue acts to build mental models of unknown domains
- Future work: engineering proactive search agents for structured knowledge bases that quickly familiarize users with an explored domain

Data repository with annotated dialogue corpus



github.com/sebischair/conversational-domain-exploration-data