Vocalizing Large Time Series Efficiently

Immanuel Trummer, Mark Bryan, Ramya Narasimha

Motivation

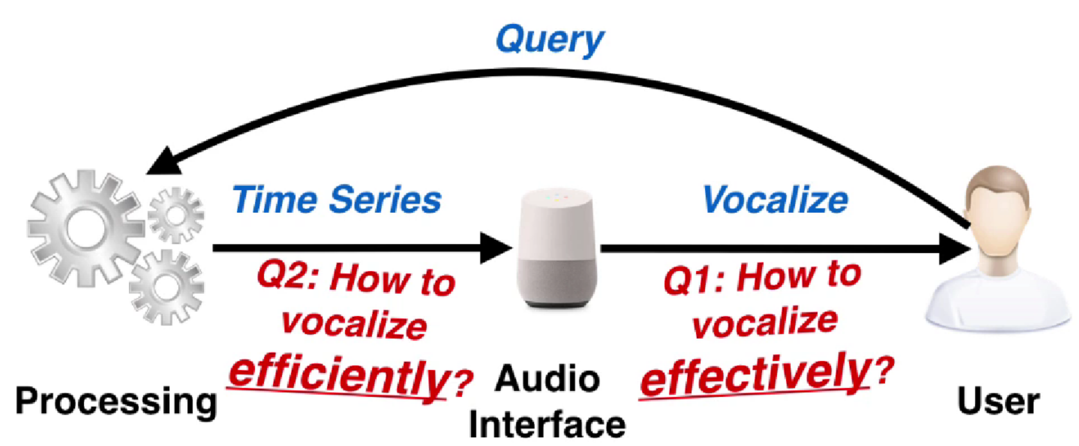

Fig. 1: Time Series Vocalization

Example Scenarios:

- Stock-market, weather trends
- Annual sales, monthly subscribers

Why Audio Interface?

- Ubiquitous nature of Alexa, Siri, Cortana
- Specially-abled people, who have no alternative

Challenges:

- Information appears gradually as opposed to at once
- System decides pace v/s you choose your focus area

Problems to solve -

- Generate answers efficiently
- Efficiency of speech planning
- Efficiency of query processing (on the fly)

Patterns & Pattern Library

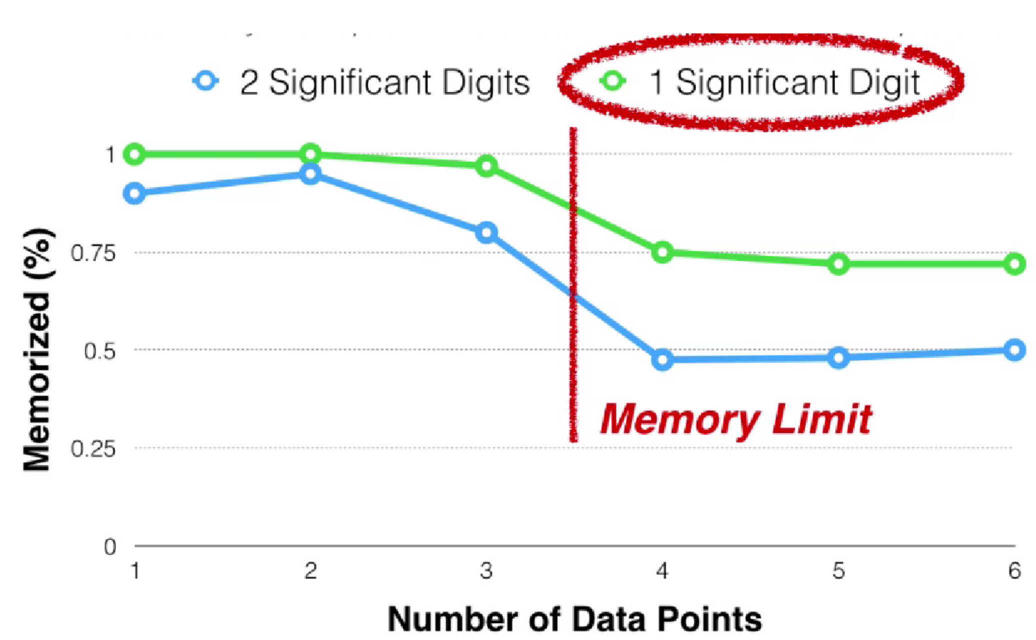

Fig. 2: Memory Limit [4, 1]

Humans can only remember 3-4 Data points with 1 significant digits. Thus, this necessitates the need for *patterns*.

Patterns				
Text Template	rises from y_1 to y_2.	falls from y_1 to y_2	remains at y	spikes to y_{max}

Template Example & Constraints

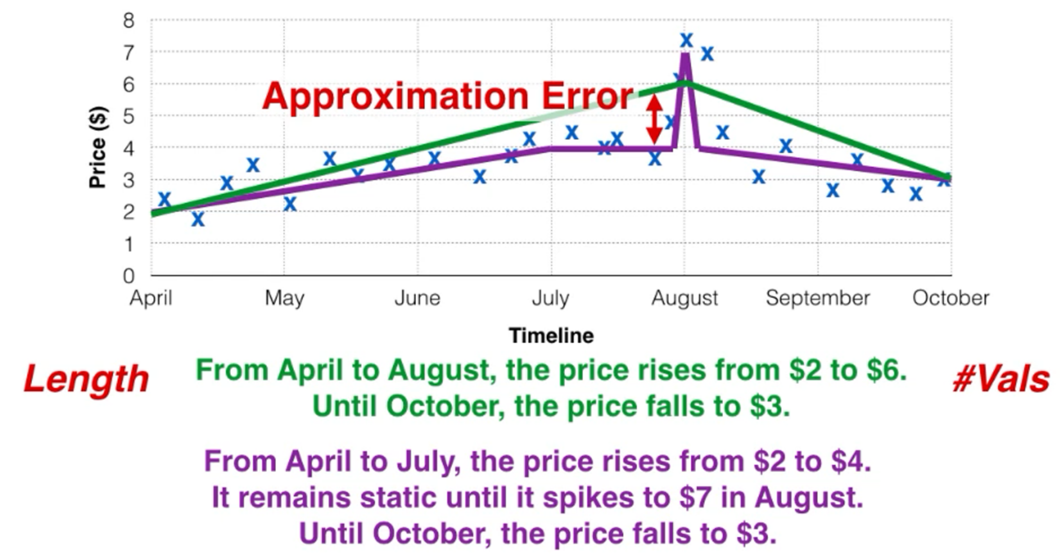

Fig. 3: Example - Price Series

Problem Statement

Given -

- *Query* producing time series data
- Library of *Speech Patterns*

Find -

- Speech minimizing *Approximation Error*
- Respecting *Length* and *Complexity* constraints

Full Query Evaluation v/s Result Sampling

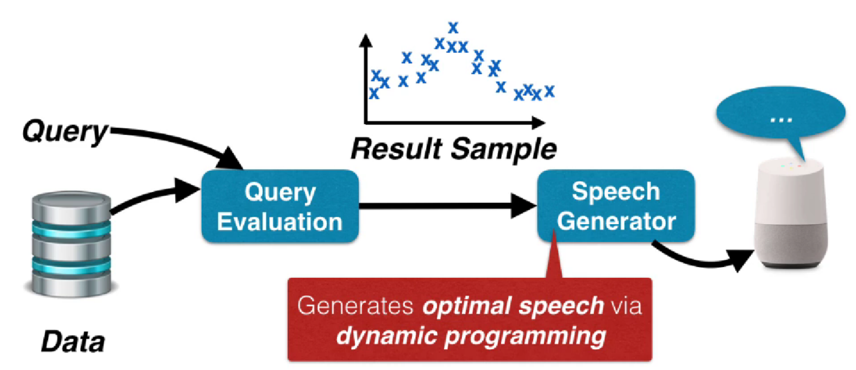

Fig. 4: Full Query Evaluation

This method gives near perfect speech, but can be computationally expensive. As this is to be implemented online, sampling can be a viable strategy.

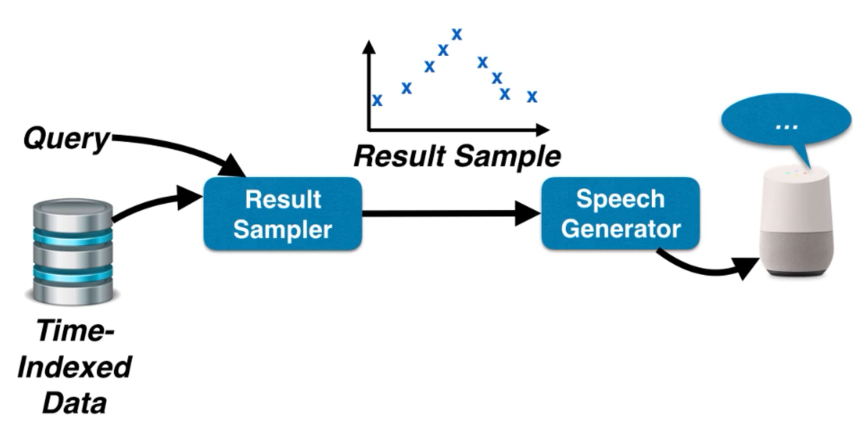

Fig. 5: Result Sampling

Formalization via Optimal Experiment Design

Optimal Experiment Design is aimed at optimizing experimental setup in order to gain the maximum amount of information. Our setup -

- **Experiment** - Result Sampling
- **Parameters** - Time points for sampling
- **Goal** - Narrow down speech choices (coarse)

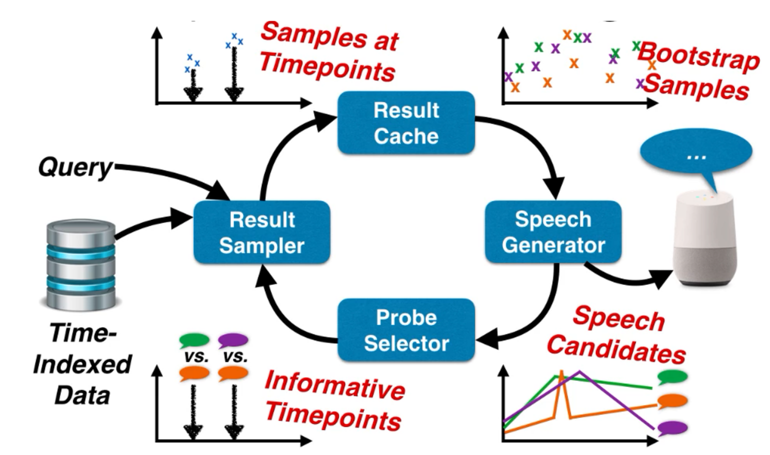

Fig. 6: Experimental Design

Initially, random samples are generated. Then, the following steps are performed iteratively - (1) Generate several *Bootstrap Samples* and the corresponding optimal speech candidate. (2) Find optimal time points where speech candidates disagree the most. (3) Feed information to *Result Sampler* to generate more points around these regions.

Once a timeout ($\approx$ 500 ms) is reached, choose the optimal speech candidate and vocalize it.

Results - Vocalization Methods

Experimental Design provides the best trade-off b/w computation time and accuracy.

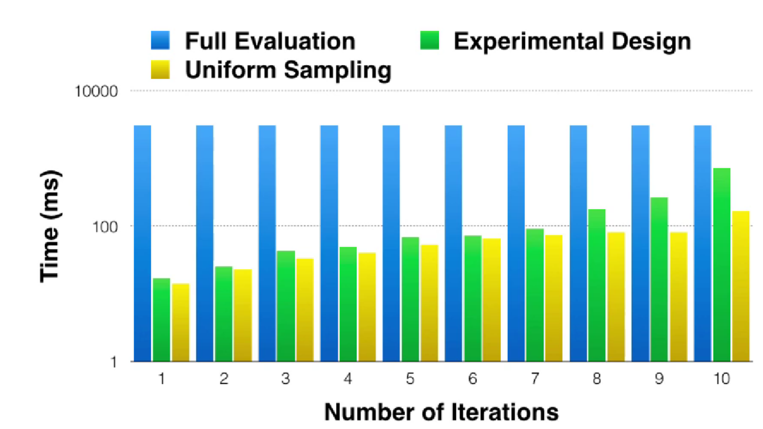

Fig. 7: Vocalization Time

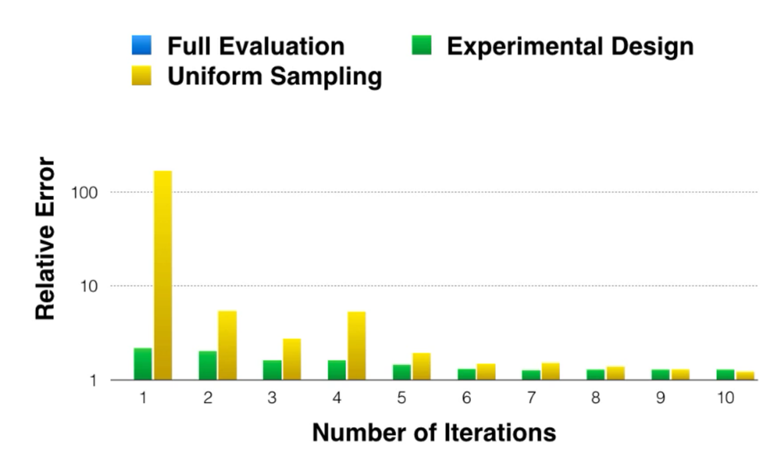

Fig. 8: Vocalization Error

Results - User Interfaces

User Study Group -

- **Participants** - 20 crowd workers (AMT)
- **Task** - propose buy/sell dates
- **Performance Metric** - Gain in USD

Criterion	Visual Interface	Audio Interface	Difference
Gain ($)	6,002	5,051	-16%
Time (s)	278	335	+21%
# Queries	8	12	+50%

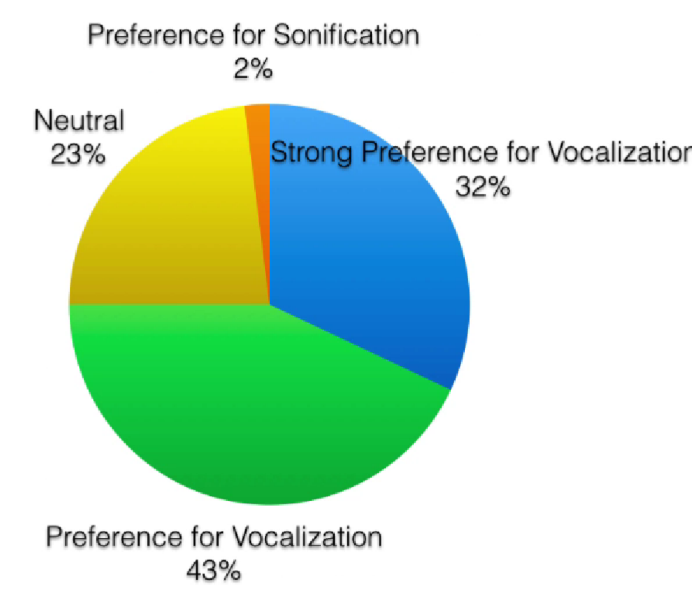

Fig. 9: Sonification [2, 3] v/s Vocalization

Summary

- Complement visualization with *vocalization*
- Approaches for *processing* and *vocalization*
- Simple data analysis possible via vocal interface
- Users prefer *speech* over non-speech output

References

[1] David Caplan and Gloria S. Waters. "Verbal working memory and sentence comprehension". In: *Behavioral and Brain Sciences* 22.1 (1999), pp. 77–94. DOI: 10.1017/S0140525X99001788.

[2] Thomas Hermann, Andy Hunt, and John G. Neuhoff. *The Sonification Handbook*. Berlin: Logos Verlag, 2011.

[3] Rameshsharma Ramloll et al. "Using Non-speech Sounds to Improve Access to 2D Tabular Numerical Information for Visually Impaired Users". In: (Jan. 2001). DOI: 10.1007/978-1-4471-0353-0_32.

[4] T.L. Saaty and M.S. Ozdemir. "Why the magic number seven plus or minus two". In: *Mathematical and Computer Modelling* 38.3 (2003), pp. 233–244. ISSN: 0895-7177. DOI: https://doi.org/10.1016/S0895-7177(03)90083-5. URL: https://www.sciencedirect.com/science/article/pii/S0895717703900835.

[5] Immanuel Trummer, Jiancheng Zhu, and Mark Bryan. "Data vocalization: optimizing voice output of relational data". In: *Proceedings of the VLDB Endowment* 10 (Aug. 2017), pp. 1574–1585. DOI: 10.14778/3137628.3137663.