

Quantum Natural Language Processing

Kick-off Meeting
14th January 2019

ICHEC, Dublin, Ireland

Agenda

- Project overview
- People at ICHEC and Intel
- Execution plan
 - Work plan and timeline
 - Risk mitigation
 - Dissemination strategy
 - Logistics
 - Action points

Project overview

- DisCo model
 - Distributional compositional semantics
 - NLP algorithm to compute meanings of two sentences and decide if their meanings match
 - Incorporates grammatical structure of sentences in a language into the analysis algorithms
 - “Closest vector problem” algorithm
 - Determine the word/phrase out of a set of words/phrases that has the closest relation to a given word/phrase
 - “CSC sentence similarity” algorithm
 - Compute the meaning of sentences/phrases and perform sentence similarity calculations
- Objectives
 - Deploy Intel Quantum Simulator (qHiPSTER) on Kay
 - Implement DisCo model algorithms on qHiPSTER
 - Evaluate performance of DisCo model algorithms
 - NOTE: Proposal does not include comparison with a classic HPC implementation
 - Period: 14-Jan-2019 to 13-Mar-2020

People at ICHEC and Intel

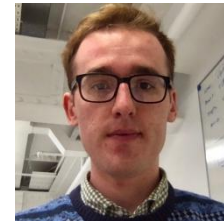
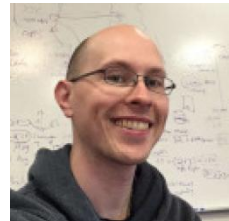
- ICHEC

- Researchers

- Lee James O’Riordan
 - Myles Doyle
 - Venkatesh Kannan (PI)

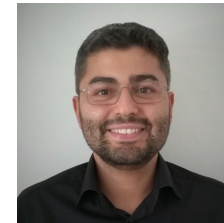
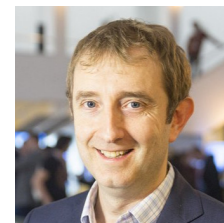
- Project manager

- Peter Woods



- Intel

- Jim Kenneally (Lead)
 - Brian Quinn
 - Fabio Baruffa



Agenda

- Project overview
- People at ICHEC and Intel
- Execution plan
 - Work plan and timeline
 - Risk mitigation
 - Dissemination strategy
 - Logistics
 - Action points

Work plan and timeline ... (1)

- Project executed through 4 tasks
 - Task 1 – Resource setup and problem mapping
 - Task 2 – Solution development
 - Task 3 – Solution evaluation and packaging
 - Task 4 – Technical report and handover
- Meetings

Month	Location
M01 (Jan 2019)	Project kick-off meeting (ICHEC, Dublin, Ireland)
M01 – M14	One meeting every 3 weeks (video conference and at Intel Ireland Campus, Leixlip, Ireland as needed)
M11 (Nov 2019)	Face-to-face meeting between ICHEC and Intel Quantum Computing team at SC 2019
M14 (Mar 2020)	Project closing meeting (Intel Ireland Campus, Leixlip, Ireland)

Work plan and timeline ... (2)

- Task 1 – Resource setup and problem mapping (M01 – M03)
 - Objective
 - Install and test the Intel Quantum Simulator (qHiPSTER) on the ICHEC supercomputer (Kay)
 - Define the strategy to map the quantum version of the DisCo model algorithm on the Intel Quantum Simulator.
 - Deliverables
 - D1.1 (M03): A report describing the abstraction of the quantum DisCo model algorithms for implementation on any quantum platform, and their mapping on qHiPSTER, along with testing and evaluation methodologies, and the representative corpora.
 - Milestones
 - M1.1 (M01): qHiPSTER is installed and tested on ICHEC's Kay.
 - M1.2 (M03): Abstraction and mapping of DisCo algorithms on qHiPSTER is complete.
 - M1.3 (M03): Testing and evaluation methodologies are defined.
 - M1.4 (M03): Representative corpora for testing are ready.

Work plan and timeline ... (3)

- Task 2 – Solution development (M04 – M11)
 - Objective
 - Implement the DisCo module algorithms on qHiPSTER and perform preliminary evaluations.
 - Deliverables
 - D2.1 (M07): A report summarising the implementation, testing and evaluation of the “closest vector problem” algorithm using the representative corpora.
 - D2.2 (M11): A technical presentation at an Intel event at SC 2019.
 - D2.3 (M11): A report summarising the implementation, testing and evaluation of the “CSC sentence similarity” algorithm using the representative corpora.
 - Milestones
 - M2.1 (M07): Implementation of the abstract and qHiPSTER-specific code, and evaluation for the “closest vector problem” is complete.
 - M2.2 (M11): Implementation of the abstract and qHiPSTER-specific code, and evaluation for the “CSC sentence similarity” algorithm is complete.

Work plan and timeline ... (4)

- Task 3 – Solution evaluation and packaging (M12 – M14)
 - Objective
 - Perform extended evaluations and package the software solution along with results and documentation.
 - Deliverables
 - D3.1 (M14): A report describing the larger corpora, and summarising the final evaluations using the larger corpora.
 - D3.2 (M14): Final version of the abstract interface code for the DisCo algorithms, with documentation.
 - D3.3 (M14): Final version of the qHiPSTER-specific code for the DisCo algorithms, with documentation.
 - D3.4 (M14): A technical poster and presentation for ISC 2020.
 - Milestones
 - M3.1 (M14): Larger corpora for final testing and evaluation is ready.
 - M3.2 (M14): Final version of the DisCo algorithms implementations and evaluation is complete.

Work plan and timeline ... (5)

- Task 4 – Technical report and handover (M14)
 - Objective
 - Delivery of the software solutions and results to Intel.
 - Deliverables
 - D4.1 (M14): A consolidation of reports from deliverables D1.1, D2.1, D2.2, D2.3, D3.1, D3.3 and final software from deliverables D3.2 and D3.3.
 - Milestones
 - M4.1 (M14): Consolidated reports and final software are prepared.
 - M4.2 (M14): Deliverable D4.1 is published on a GitHub page under the Apache License version 2.0.

Risk mitigation

Risks	Steps for management
Installation and performance issues of the Intel Quantum Simulator (qHiPSTER) on the ICHEC Kay supercomputer.	The first task in the project work plan is to install and test qHiPSTER on the ICHEC supercomputer. Reservations can be made for this project on the Kay supercomputer to achieve the desired performance. The simulator developers from Intel will also be available to assist with solving installation and configuration issues.
Issues in mapping and/or implementing the DisCo model algorithms on the programming model of qHiPSTER.	The developers of qHiPSTER will be available to provide user support during implementation of the algorithms. Also, the authors of the DisCo model at the University of Oxford may be consulted for a deeper understanding of their algorithms.
Performance of the algorithm implementations is lower than expected and thus evaluation is delayed.	The size of the corpuses targeted for the evaluations may be reduced accordingly to achieve reasonable performance to get evaluation results.

- Intel support is primarily sought for troubleshooting for installing and using the qHiPSTER platform.
- What is the channel for ICHEC to raise a question or issue with Intel?

Dissemination strategy

- Proposed
 - SC 2019
 - ISC 2020
- Any others
 - Press release
 - 2nd All Ireland Conference on Quantum Technologies at Maynooth University, 21-Jan-2019
 - EQTC 2020 (TBC)
 - Intel Technical Computing Roundtable 2020
 - Talk at a quantum computing workshop at UCD (TBC), Mar-2019
 - ...

Logistics

- ICHEC-Intel periodic meetings (proposed)
 - Once every 3 weeks
 - Mondays 11:00 (Irish time) starting 04-Feb-2019
 - Google meetings at <https://meet.google.com/zvp-kmhp-pak>
- Hosting and delivery of code and reports
 - <https://git.ichec.ie/intel-qnlp/intel-qnlp/>
 - Currently public. Access restrictions?
- Any other business?

Action points

- Start technical tasks for M01 (ICHEC – VK, LOR, MD)
 - T1.1. Install and test the Intel Quantum Simulator (qHiPSTER) on Kay
 - T1.2. Investigate the DisCo model algorithms and define their mapping strategy on qHiPSTER
- Set up periodic meetings on calendar (ICHEC – VK)
 - ICHEC-Intel meetings
 - ICHEC-internal meetings
- Upcoming milestones
 - M1.1 (M01): qHiPSTER is installed and tested on ICHEC's Kay.