A Quantum Annealing Approach to Biclustering

L. Bottarelli, M. Bicego, M. Denitto, A. Di Pierro, A. Farinelli
Università di Verona

12-13 December, 2016 – TPNC 2016, Sendai (Japan)

Table of contents

Introduction

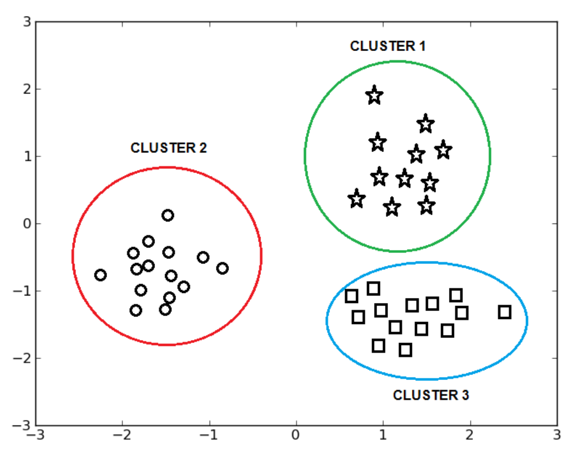
Motivation

- Problems in Artificial Intelligence are typically intractable.
- Heuristics, greedy approaches etc are used instead of exact algorithms.
- Can the D-Wave quantum annealing hardware be helpful?
- No final answer at this stage.
- We investigate this for the biclustering problem

Biclustering

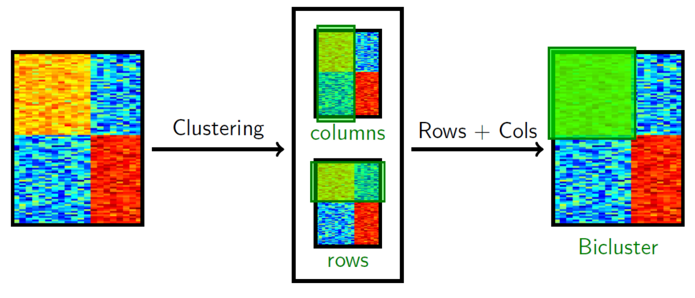
Clustering

Partition into groups a set of objects based on the concept of similarity



Biclustering

Perform **simultaneous** clustering of rows and columns of a given data-matrix

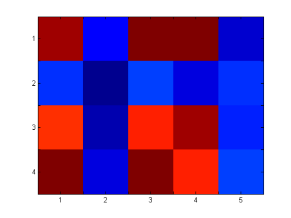


Biclustering

- Used in different scenarios, the most important is **expression microarray data analysis**
  - Rows and columns represent genes and experiments respectively
  - Biclustering can provide invaluable information to biologist
- We focus on the **One-Bicluster** problem
  - Standard technique where biclusters are extracted one by one
  - We extract the largest bicluster
  - The other biclusters can be retrieved by masking the previous ones

Biclustering

- **Input:** real-valued matrix. Each entry $a_{i,j}$ represents the activation level of gene i under the experimental condition j



- **Output:** binary matrix that represents which elements belong to the bicluster

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

- Bicluster elements must comply a **topological constraint**

Valid configuration

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

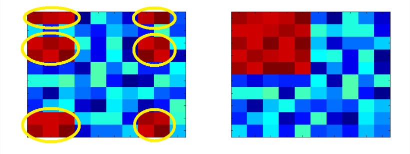
Invalid configuration

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

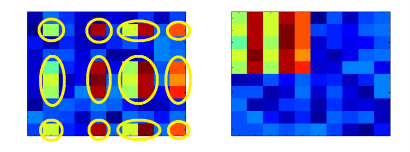
Biclustering

- The similarity between elements is defined by a coherence measure
- We used two types of coherence:

Constant coherence $O_{i,j,t,k} = w|a_{i,j} - a_{t,k}|$



Additive coherence $O_{i,j,t,k} = w(a_{i,j} + a_{t,k} - a_{t,j} - a_{i,k})^2$



Quantum Annealing

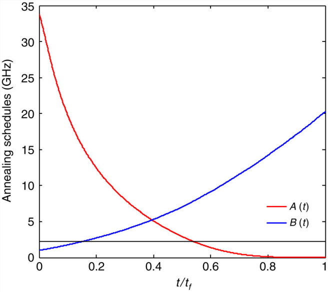
Quantum Annealing

- Conjectured in 1984 as an improvement of simulated annealing
- QA is typically achieved by the Ising Hamiltonian:

$$H(t) = A(t)H_X + B(t)H_P, \;\; t \in [0, t_f].$$

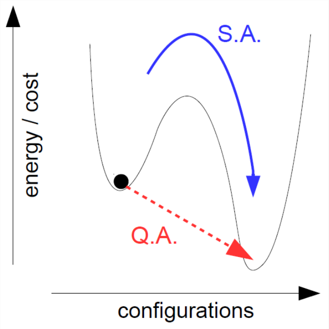
- The solution to an optimisation problem is encoded in the ground state of H_P:

$$H_P = \sum_{i \in \mathcal{V}} h_i \sigma_i^z + \sum_{(i,j) \in \mathcal{E}} J_{ij} \sigma_i^z \sigma_j^z$$



Quantum Tunneling

- Quantum fluctuations can help "tunneling" through barriers
- Numerical and experimental results have suggested that quantum tunnelling can be more effective than thermal fluctuation for reaching the ground state.
- This is still the subject of active scrutiny



QUBO

Quadratic Unconstrained Binary Optimization (QUBO)

- Often it is more convenient to model with 0/1-valued variables than the $1/-1$-valued Ising variables.
- This results in an equivalent formulation of the Ising energy minimisation problem as a QUBO problem.
- QUBOs are Ising models where the spin variables $s_i \in \{+1, -1\}$ are transformed to binary-valued variables $x_i \in \{0, 1\}$.
- This transformation is easily realised through $s_i = 2x_i - 1$.

Quadratic Unconstrained Binary Optimization (QUBO)

Goal

Find the assignment to a set of binary variables $x_1 \ldots x_n$ so as to minimize the function:

$$O(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i + \sum_{1 \leq i < j \leq n} b_{i,j} x_i x_j$$

Every instance of a QUBO can be represented by a graph

Example

$$O(x_1, x_2) = 2x_1 - x_2 + x_1 x_2$$

QUBO Model for One-Bicluster problem

Objective function:

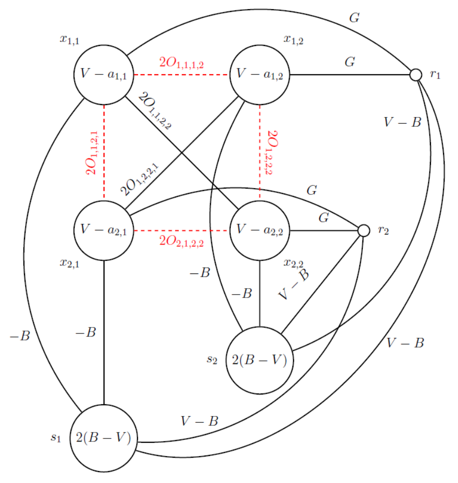
$$\arg\max_{(c_{1,1},\cdots,c_{N,M})}\left(\sum_{i,j} a_{i,j}c_{i,j} - \sum_{i,j,t,k} O_{i,j,t,k}c_{i,j}c_{t,k} + \sum_{i<t} B_{i,t}\right)$$

Complete model for a 2x2 data-matrix

$N \times M$ matrix

$\Downarrow$

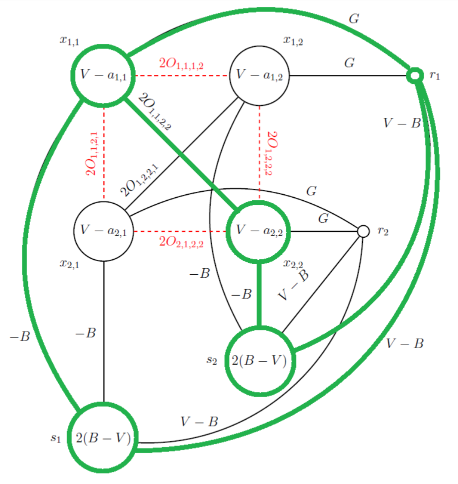
$N \times M + N + M$ nodes



QUBO Model for One-Bicluster problem

Non-valid configuration

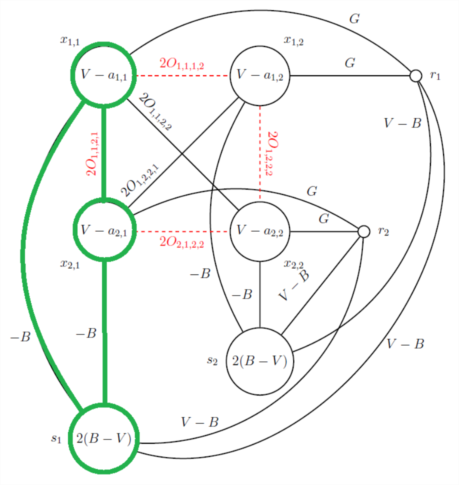
$$\sum = G > 0$$



QUBO Model for One-Bicluster problem

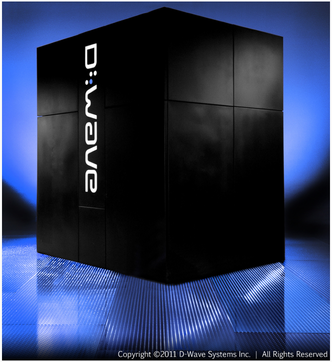
Valid configuration
$\sum = 0$

Model validity
For all valid assignments, the extra constraints provide a null contribution to the objective function. For all other configurations the contribution is > 0.



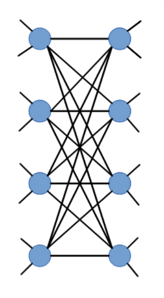
D-Wave

- First commercially available quantum annealer
- Two have been sold:
    - Lockheed Martin
    - Google - NASA
- Finds the ground state of **Ising spin glass models** or equivalently solves **Quadratic Unconstrained Binary Optimization** (QUBO) problems



Minor Embedding

- The D-Wave has a well defined **fixed architecture**
- Create a **logical model** and map it into the D-Wave architecture
- Chimera is not a **complete graph**
- A single logical qubit may need more than one **physical qubit** to be represented
- The **minor embedding** technique finds the sets of physical nodes that represent the logical ones



Minor Embedding

We performed the minor embedding using the **official D-Wave libraries** provided by J. Cai, B. Macready and A. Roy

Results of our experimental analysis

	(D-Wave 2)	(D-Wave 2X)	(Future ?)
Qubits	**512**	**1152**	**2048**
Max data-matrix size	5×6	7×7	8×8

The model proposed is very accurate but has a high space complexity:

- $\mathcal{O}(n^2 m^2)$ edges (quadratic terms)

Suitability of Quantum Annealing

Criteria for problems that are suitable for treatment with a quantum annealer:

Landscape Analysis

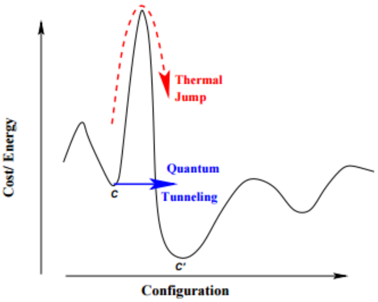
Solutions to the problem are valuable or interesting.

Runtime Advantages

Must exhibit a significant speed-up compared to other (classical) algorithms used for that problem.

Landscape Analysis

- The energy landscape must have tall and narrow barriers
- Frustration of the system implies that the energy landascape becomes extremely rugged (tall and narrow barriers)
- The biclustering Hamiltonian (both logic and embedded) gives rise to a frustrated system.



Quantum Speed-up?

- Based on the current results, one cannot claim a quantum speedup for D-Wave 2X, as this would require that the quantum processor in question outperforms the best known classical algorithm.

Quantum Speed-up?

- Based on the current results, one cannot claim a quantum speedup for D-Wave 2X, as this would require that the quantum processor in question outperforms the best known classical algorithm.
- E.g. the Hamzede Freitas-Selby algorithm

Quantum Speed-up?

- Based on the current results, one cannot claim a quantum speedup for D-Wave 2X, as this would require that the quantum processor in question outperforms the best known classical algorithm.
- E.g. the Hamzede Freitas-Selby algorithm
- However, the D-Wave team believe that such solvers will become ineffective for the next generation of annealers currently being designed. The primary motivation for this optimism is that the connectivity graphs of future annealers will have higher degree.

Quantum Speed-up?

- Based on the current results, one cannot claim a quantum speedup for D-Wave 2X, as this would require that the quantum processor in question outperforms the best known classical algorithm.
- E.g. the Hamzede Freitas-Selby algorithm
- However, the D-Wave team believe that such solvers will become ineffective for the next generation of annealers currently being designed. The primary motivation for this optimism is that the connectivity graphs of future annealers will have higher degree.

Conclusion

Conclusions and Future Work

Conclusions

- The model proposed allows us to solve the biclustering problem with high accuracy
- Space complexity of the model + current D-Wave architecture do not allow to address big matrices

Future Work

- Study new techniques in order to decompose big instances into smaller ones