

ESF Exploratory Workshop on

Correlations In Computer Science

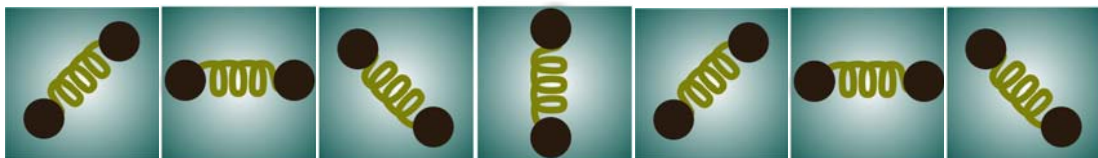
Vielsalm (Belgium), 5-8 August 2008

Convened by Ellie D'Hondt, Vrije Universiteit Brussel, Department of Computer Science

SCIENTIFIC REPORT

Report on the ESF Exploratory Work- shop on Correla- tions in Computer Science

ELLIE D'HONDT



Executive summary

The goal of the Correlations in Computer Science program is to establish the base requirements of a programming paradigm based on correlations. Correlations are mutual relationships or connections between two or more concepts, and abound in computer science. However, we do not currently know how to program with correlations. Indeed, they are usually considered a burden in application development, reusability and evolution, and are not well captured in any programming paradigm. What we propose here is to turn things upside down and to provide correlations as first-class objects in a programming environment. To do this we need a stock of examples, from which we can identify the necessary concepts and design the basic building blocks of the paradigm. We focused on three emerging areas of computer science in particular: aspect-oriented software development (AOSD), bio-inspired computing (BC) and quantum computing (QC). These cover a whole spectrum of situations, going from correlations in high- to low-level programming languages (AOSD vs. QC), and between multi as well as many objects (BC vs. others).

Since there was no backbone of research to rely on and attendees came from varied disciplines, it was crucial to set up a common vocabulary from the very start. For this reason a website incorporating a discussion group was set up (www.correlationspace.org), and a 3-page manifesto was written

entitled 'What is a correlation?'. The latter contained a first definition of the concept of a correlation and established guidelines for the preparation of talks. In particular, participants were encouraged to attack this definition and come up with at least one concrete, kindergarten example of a correlation in their research. The idea was that this would allow us to start from a concrete set of examples, smoothing the way for a quick-start of brainstorming sessions.

A second means of facilitating interactions between participants was the format of the workshop. Though people within disciplines had typically met before, there were few connections between disciplines. To avoid clustering within areas of expertise - which would have been catastrophic for the success of the workshop - we aimed at providing an optimal environment for interaction. Hence we decided upon an informal, community-forming approach, in which all participants stayed in the same venue for the duration of the workshop. The chosen location proved ideal: a former hotel providing private rooms for all participants, but also a large communal area for work and relaxation. The opening reception ensured smooth introductions, while a poster with pictures of all participants was put up as backup for identification. All meals were provided sur place in a cosy environment, and the sauna proved to be an excellent place for bonding. Though seemingly peripheral, this framework of interactivity contributed largely to the success of the workshop, in that it enabled the quick establishment of a community with common research interests. Interest was nurtured, communication facilitated, motivation increased, willingness to pursue this topic further heightened.

Scientific interaction required first and foremost that participants were introduced to each other's area of discipline and the particular research they conducted therein, keeping in mind the common framework of correlations. While the latter was highlighted through a keynote talk, the former was achieved by splitting up the first day into three sessions, respectively grouping talks on AOSD, BC and QC. Achieving a common basis of knowledge in such a short time proved to be rather ambitious. In the AOSD session in particular, even at the level of the introductory talk, there were many questions and some discussion was required to have all participants understand the basics of the field and the problems it tries to tackle. This caused some delays in the schedule, as a consequence of which part of the QC session had to be rescheduled for the next day.

The program proved to be quite dense not only because many short talks had to fit into one day but also because participants found it difficult to adhere to the proposed guidelines for talks. As mentioned above, it was suggested beforehand that attendees attack a first definition of correlations and explain one kindergarten example thereof within their half-hour talks. While some participants did try to follow this procedure, others instead gave an overview of their work. Even in the former case it proved hard to focus on a concrete example without a common framework in mind and without sharing the research context with the audience. This meant that more material had to be

explained and digested in a short time than foreseen. Two reasons were identified for the unrollment of the program as such. The first is that some people, being new to the group as a whole, felt they should present their body of work and its place within the domain context, rather than picking out one concrete example. The second, more important reason was that it was just too soon for participants to be able to follow up on the manifesto guidelines. In effect, participants first needed a more concrete understanding of what correlations are and how they appear in different guises across these domains. This was certainly achieved at the meeting, and at the same time people did build up a common vocabulary, greatly facilitating the rest of the workshop and, indeed, the continuation of this research program. At the same time general principles and example correlations were touched upon, forming the root of discussions later on in the workshop.

The second day of the workshop was more oriented towards developing tools and techniques for building up a correlation paradigm, mediated through another keynote talk and a brainstorming session. General principles, alternative definitions of correlations, and candidate examples were collected throughout talks and an overview of these was presented as an introduction to the brainstorming session. Because the main goal of research talks proved to be the establishment of a common vocabulary rather than executing the proposed manifesto, we adapted the planning dynamically so that the latter formed the basis of the brainstorming sessions. As such predefined groups of four pondered the main questions posed (what are the basic building blocks of a correlation paradigm? what are good examples?) and presented their results to the whole group afterwards. Brainstorming sessions were enormously successful in that all groups managed fluent communication, which was absolutely not trivial considering the maximal variation of skills within each group and the fact that group members had not collaborated previously. On top of this, common notions emerged from all groups, which strengthens our case for a correlation paradigm and ensures that we are on the right track for further development of this research.

The workshop concluded in style with an excellent dinner and lots of discussion. All participants were extremely enthusiastic about the workshop, which was generally considered a success. A consensus view was that another workshop is required to further establish the research topic as well as the community. Initiating head-on concrete research in this new interdisciplinary topic -- by forcing people to come up with working examples -- proved to be difficult without first establishing some common ground in this newly established community. Within the boundaries of a two-day workshop, we nevertheless achieved the common insight that this is valuable and necessary research, and the willingness to pursue the development of a correlation paradigm further.

Scientific content

The two-day workshop on focused on three different areas as foraging grounds for correlations in computer science: aspect-oriented software development (AOSD), bio-inspired computing (BC) and quantum computing (QC). On the other hand it was crucial to set up communication channels across disciplines. For this reason part of the workshop was dedicated to talks by participants, split up accordingly in a session each on AOSD, BC and QC, while an equal part was devoted to brainstorming sessions in smaller groups. This was complemented by two longer keynote talks by senior researchers, wherein the speakers had the liberty to broaden and deepen their presentation topic. We refer to the section 'Final Programme' for details on the exact planning of the workshop.

The scientific part of the workshop opened with a first keynote talk on 'Correlations in computer science: a partial overview' by Samson Abramsky, The idea was to have this renowned and experienced computer scientist indicate a number of situations in computer science where correlations appear. After placing the notion of correlation within the broad context of computer science and presenting some general principles, three main correlation stories were laid out: "Information dependence and independence" (relational databases), "Synchronization and interaction" (processes), and "The logic of quantum information flow" (quantum entanglement). The talk relied heavily on logic and moreover promoted the latter as a technique for developing a theoretical framework for correlations.

After the keynote talk three sessions followed with shorter, half-hour talks.

AOSD session. As the jargon from the AOSD community was new to most participants not in this session, a general introduction to AOSD was first given by Michael Haupt. This was followed by a number of research talks, of which some, ironically, advocated the need for new techniques to deal with aspects. Indeed, Michael Haupt spoke about an aspect machine model in terms of delegate functions, where aspects become first-class semantical objects of the system. Also Jim Coplien spoke about dealing with aspects within the DCI paradigm, where Data modeling, Context and Interaction are three interacting parts of the design that embody different types of correlations. The other talks dealt more with the notion of correlations within this field. Rémi Douence argued that correlations are always a consequence of the chosen decomposition mechanism, and that multiple, alternative decomposition mechanisms combined with functional mappings between them are needed to capture correlations. Theo D'Hondt proposed a key example of correlations in software - providing recursion support in a basic interpreter - and formulated some general principles and questions.

Bio-inspired computing session. Since the topics covered in this session proved to be dealing with disparate subjects, no introductory talk was included in the program. Speakers instead covered a number of situations that can be viewed as correlation examples and gave techniques to model these. Devdatt Dubhashi proposed weak linkage of genes as a biological correlation, one that is adaptive, indirect and logical (through gene regulation changes rather than changes of genes themselves). Vincent Danos discussed a process algebraic language for analysing cellular signalling pathways, the latter being the envisaged correlations. The joint talk given by Olivier Michel and Jean-Louis Giavitto sketched two notions of correlation that appear when one implements simulation models of complex dynamical systems: one related to interaction and the other to the global description of system dynamics. In both cases, the correlation is between the interacting, local part and the evolution of the system. Finally, Radu Mardare discussed an algebraic approach to the formalization of the concept of information for biomolecular systems.

Quantum computing session. This session was again opened with an introductory talk by Damian Markham, to ensure all participants had the basics under their belt. The same speaker continued with a brief outline of the appearance of entanglement in physics, and how it is useful for quantum information. This was followed by a survey of the measurement-based model of quantum computation, and in particular the measurement calculus, by Elham Kashefi. These two talks nicely complemented each other, as the measurement-based model relies on a particular type of entanglement, namely graph states, for its execution. The next two talks were shifted to the next day due to the delayed schedule mentioned in the above. Though they were preceded by another keynote talk, we discuss them here as they were part of this session. A second joint talk, this time by Sonja Smets and Alexandru Baltag, covered the dynamic-epistemic logic approach to quantum computation, a body of work almost entirely set up by the presenters. In particular this formalism was used to analyse quantum entanglement and the information flow it induces. Finally, the convenor closed the quantum session with a review of entanglement programs, i.e. computations in which entanglement is used directly and concretely to steer quantum computations. Roughly three areas were identified: entanglement as a function; entanglement as a distributed primitive, and entanglement as data representation.

The second day of the workshop was opened with another keynote talk by Karl Lieberherr. Again an experienced researcher was attracted to discuss his ideas on possible techniques for constructing a correlation paradigm. His concrete proposal is reflected in the title of the talk: 'Separation of Concerns in Algorithmic Trading Robots: Towards a Correlation Paradigm'. Trading robots are considered a lead example, both applicable

to economy and biological contexts. The proposed technology is DemeterF, a generic technology that combines ideas from adaptive programming, functional programming, and datatype-generic programming to address traversal related concerns. It allows to capture and separate the essential aspects of functional traversals from the structures to be traversed, the former being the correlations between the latter.

The rest of workshop was devoted to a brainstorming session. Participants were divided into four predefined groups, such that each group had a varied skill set and limited opportunity for subgrouping in that people were new to each other. A short introduction was given by the convenor, in which correlation-related statements gathered during the meeting were presented. Also each group was presented with the same task: defining the basic ingredients of a paradigm for problem-solving with correlations, which is tailorable to the different contexts encountered during the workshop. Groups split up for the afternoon and presented their results in a final group session, which was also used to discuss future planning.

Assessment of results

As one of the referees correctly noted the topic of this workshop is very 'blue sky'. We are embarking on an ambitious project of having people from different disciplines devise a new way of programming based on the notion of a correlation. One could say the main contribution of this whole endeavour was the conception of the notion of a correlation as a computational concept in itself, and the idea of building up a programming paradigm from it. While this proposal was key in attracting people to the event, the workshop itself was crucial in going beyond people's curiosity. A basis was laid for a new domain within computer science, interdisciplinary because it affects paradigms appearing in new (bio-inspired and quantum) as well as standard (software engineering, aspects) models of computing. All participants left convinced that correlations exist and that the similarity in correlation examples across disciplines is real rather than metaphorical. This core group agreed in that creating a correlation paradigm is a necessary, useful, and worthwhile research program. We stress that in starting up this domain we had nothing much to go on. Of course we can look at examples of paradigm building in classical computer science and tools and techniques used therein. But next to that, we are on our own: correlations are rather counterintuitive objects, and some of the example areas are still struggling to define themselves in terms of computation. However, our future plan is precisely to aid in this process of paradigm building by working with this new notion of correlation and cross-fertilising results across fields through this general framework of a correlation paradigm.

The short-term goal of this workshop was to come up with a formal definition of a correlation that encompasses all examples presented at the workshop. To quick-start this development a

document was shared with all participants well before the event took place, presenting a first proposal for a correlation definition, as follows.

A correlation is the encapsulation of a concern that cannot be separated under a particular decomposition mechanism, i.e. an entity capturing the behaviour and functionality of a crosscutting concern.

Participants were encouraged to present a concrete, kindergarten example of a correlation within their field within their half-hour research talks. This approach was a bit too optimistic, in that one cannot expect these issues to be clear to all beforehand. Indeed the workshop itself turned out to be the forum to converge on a definition and figure out what constitutes an adequate correlation example. Hence, while some people proposed alternative definitions, general guidelines, and possible examples, a significant part of the workshop was taken up by laying the foundation for communication between participants to be possible at all. Indeed, we needed to establish a common vocabulary before being able to decide on the similarities between correlation situations. The first day of research talks together with lots of possibilities for discussion were successful in achieving precisely this. This resulted in effective brainstorming sessions: indeed communication between these heterogeneous groups of four would not have been possible without a prior sharing of basic knowledge. Considering the short period of time and the fact that participants were new to each other we feel that this is an extremely positive outcome of the workshop.

The brainstorming sessions turned up a number of issues very useful for the formal development of a correlation framework. In fact a number of these emerged independently in several subgroups. We list the main arguments here.

1. Correlations appear when one moves between global and local descriptions of a system. They deal with interaction between subsystems, and are generally dependent on the particular means chosen for subdividing the whole. In fact the latter defines what is considered local in a system. QC is different to the other areas in that here any chosen separation of concerns leads to correlations. Correlations together with local information lead to global information.
2. Correlations are a static representation of the dynamic interaction between parts of a system. One should make the distinction between situations dealing with construction and with execution. In the former one is concerned with how a particular system functions as a whole and how this functionality arises. Here one wants to make assertions about the system and, through a refinement function, reduce these to expressions pertaining to the system's parts. Correlations then appear as those ingredients that cannot be attached to a (local) sub-state. In the latter one starts with a sys-

tem's states together with a evolution (rather than a refinement) function, the evolution then dictating the correlations that will emerge.

3. The more correlations, the less robust a system is. But note that there is a transition in this law at some point: a highly correlated biological system built up from many building blocks (e.g. the heart) is very robust. The complexity of the correlation structure in itself may be important here. Note also that in QC there are many ways of defining an order on entangled states, and hence no unique way of determining if a state is more entangled than another.

In the near future we will construct a diagram relating the concepts that came out of the brainstorming sessions. Examples collected from the workshop will be rephrased in terms of this diagram. From this backbone a first formalised framework can be proposed, such that all examples can be rephrased in terms of it. This work will be shared through the workshop discussion group, so that participants can ameliorate and extend the framework by iteration. However, it was agreed by all participants that to really make this research advance, and the collaboration started up in this new community last, we need another workshop. We are now at a stage where all participants understand the need for a correlation paradigm and see that there is a commonality between the fields included. We converged on a definition and basic principles. People are now ready to do research on the problem, and this is what we should get together on in another year or so. Everybody expressed a keen interest in such a second event, and this is our future target. This would be no longer an exploratory workshop, but rather the usual workshop format with a prior call for papers, giving rise to actual research. In this way we build up a body of work which would, in turn, facilitate brainstorming sessions at a more concrete level, engendering new results and research possibilities. We are currently looking into possibilities for organising such a second event.

Final Programme

Tuesday 5 August 2008

- 17:00 Minibus departure from Brussels Airport/Brussels
 South Train Station
- 19:00 Arrival in Hotel Belle Vue, Vielsalm
- 20:00 - Opening dinner
22:00

Wednesday 6 August 2008

- 8:00 - *Breakfast*
9:00
- 9:00 - **Ellie D'Hondt**: Meeting introduction
9:15
- 9:15 - **Kaise Sere**: Presentation of the European Science
9:30 Foundation (ESF)
- 9:30 - **Samson Abramsky**: Correlations in computer science:
10:30 a partial overview (*Keynote talk*)
- 10:30 - *Coffee break*
11:00

Session 1: Aspects-oriented correlations

- 11:00 - **Michael Haupt**: Introduction to Aspect-Oriented
11:30 Software Development
- 11:30 - **Michael Haupt**: How Aspects Might Correlate
12:00
- 12:00 - **Rémi Douence**: The democracy of multiple decompo-
12:30 sitions.

- 12:30 - *Lunch*
14:00
- 14:00 - **Jim Coplien**: A New Object Model with Behavioural
14:30 Correlations: The other shoe drops from MVC's
creator
- 14:30 - **Theo D'Hondt**: About aspects and mental models
15:00

Session 2: Correlations in Biocomputing

- 15:00 - **Devdatt Dubhashi**: Software Architecture Principles
15:30 of Weak Linkage in Gene Regulatory Systems.
- 15:30 - **Vincent Danos**: A stochastic calculus of binding
16:00
- 16:00 - *Coffee break*
16:30
- 16:30 - **Olivier Michel & Jean-Louis Giavitto**: Space in
17:00 Correlations & Correlations in Space
- 17:00 - **Radu Mardare**: BioLogics: biomolecular information
17:30 in logical form

Session 3: Correlations in Quantum Computing

- 17:30 - **Damian Markham**: Introduction to Quantum Computa-
18:00 tion
- 18:00 - **Damian Markham**: Entanglement in Physics
18:30
- 18:30 - **Elham Kashefi**: Computing with Entanglement
19:00
- 20:00 - *Dinner & Bar discussion (identification of 2 cor-
22:00 relation problems)*

Thursday 7 August 2008

9:00 - **Karl Lieberherr**: Separation of Concerns in Algorithmic Trading Robots: Towards a Correlation Paradigm (*Keynote talk*)
10:00

(continued) Session 3: Correlations in Quantum Computing

10:00 - **Sonja Smets&Alexandru Baltag**: A Dynamic-Epistemic Perspective on Quantum Correlations
10:30

10:30 - **Ellie D'Hondt**: Programming with Entanglement
11:00

11:00 - *Coffee break*
11:30

11:30 - Introduction of correlation problem & discussion
12:30

12:30 - *Lunch*
14:00

Session 6: Solving the correlation problem

14:00 - Brainstorming in 4x4 workgroups
16:30

16:30 - Presentation of results
17:30

17:30 - Discussion session: future research and planning
18:30

20:00 - *Closing Dinner*
22:00

Friday 8 August 2008

8:30 - 9:30 *Breakfast*

10:00 Minibus departure from Hotel Belle-Vue, Vielsalm

12:00 Arrival in Brussels Airport/Brussels South Train Station

Final list of participants

Compared to the original application for this workshop some shifts occurred in the participant list. One reason was conflicting schedules, even though the dates of the workshop were communicated well in advance (Gabriel Ciobanu, Marian Gheorghe, Tom Lenaerts, Centre for Integrative Bioinformatics /Netherlands). This was in some cases nicely solved by having a replacement participant from the same lab (Stijn Mostinckx □ Theo DHondt, Mira Menzini □ Michael Haupt, Hohmann Laboratory/Sweden □ Devdatt Dubhashi from Chalmers University of Technology/Sweden). As an added plus all these replacements were in fact better suited to the context of the workshop, as all these people had a better feel for programming paradigms. A second reason was loss of interest in the event (Computational Biophysics Group/Germany, Department of Molecular Cell Physiology/Netherlands). Other people were attracted, keeping in mind insofar as possible the necessity of diversity in skills and location (Damian Markham). The resulting group was kept deliberately smaller than in the original proposal, as it was felt that this would improve interactivity between all members of the workshop.

Convenor:

1. **Ellie D'HONDT**
TINF
Vrije Universiteit Brussel
Pleinlaan 2
1050 Brussel
Belgium
eldhondt@vub.ac.be

4. **Alexandru BALTAG**
Computing Laboratory
University of Oxford
Wolfson Building
Parks Road
Oxford OX1 3QD
United Kingdom
Alexandru.Baltag@comlab.ox.ac.uk

ESF Representative:

2. **Kaisa SERE**
Department of Computer Science
Natural Sciences and Mathematics
Abo Akademi University
Lemminkäisenkatu 14
20520 Turku
Finland
kaisa.sere@abo.fi

5. **Jim COPLIEN**
Gertrud&Cope
Mørdrupvænget 14
3060 Esbjerg
Denmark
jcoplien@gmail.com

6. **Theo D'HONDT**
PROG
Vrije Universiteit Brussel
Pleinlaan 2
1050 Brussel
Belgium
tjdondt@vub.ac.be

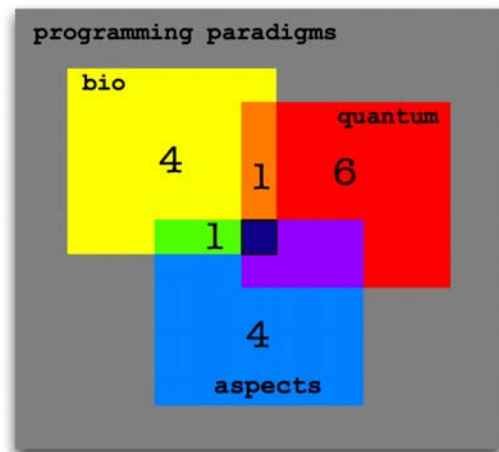
Participants:

3. **Samson ABRAMSKY**
Computing Laboratory
University of Oxford
Wolfson Building
Parks Road
Oxford OX1 3QD
United Kingdom
Samson.Abramsky@comlab.ox.ac.uk

7. **Vincent DANOS**
 School of Informatics
 University of Edinburgh
 James Clerk Maxwell Building
 Mayfield Road
 Edinburgh EH9 3JZ
 United Kingdom
vincent.danos@gmail.com
8. **Remi DOUENCE**
 Département Informatique
 Obasco Group
 École des Mines de Nantes
 4, rue Alfred Kastler
 44307 Nantes cedex 3
 France
Remi.Douence@emn.fr
9. **Devdatt DUBHASHI**
 Chalmers University of Technology
 Department of Computing Science and
 Engineering
 412 96 Göteborg
 Sweden
dubhashi@cs.chalmers.se
10. **Jean-Louis GIAVITTO**
 IBISC
 Université d'Evry Val d'Essone
 Tour Evry 2, GENOPOLE
 523, Place des Terrasses de l'Agora
 91000 Evry
 France
gjavitto@ibisc.fr
11. **Michael HAUPT**
 Software Architecture Group
 Hasso Plattner Institute for Software
 Systems Engineering
 Universität Potsdam
 Postfach 900460
 14440 Potsdam
 Germany
michael.haupt@hpi.uni-potsdam.de
12. **Elham KASHEFI**
 School of Informatics
 University of Edinburgh
 James Clerk Maxwell Building
 Mayfield Road
 Edinburgh EH9 3JZ
 United Kingdom
ekashefi@gmail.com
13. **Karl LIEBERHERR**
 College of Computer and Information
 Science
 Northeastern University
 MS WVH-202
 360 Avenue of the Arts
 Boston MA 02115-5000
 United States
lieber@ccs.neu.edu
14. **Radu MARDARE**
 Centre for Computational and Systems
 Biology
 Microsoft Research Trento
 Piazza Mancini 17
 Povo
 38100 Trento
 Italy
mardare@cosbi.eu
15. **Damian MARKHAM**
 Laboratoire PPS
 Université Paris Diderot - Paris 7
 Case 7014
 75205 Paris Cedex 13
 France
damian.markham@gmail.com
16. **Olivier MICHEL**
 IBISC
 Université d'Evry Val d'Essone
 Tour Evry 2, GENOPOLE
 523, Place des Terrasses de l'Agora
 91000 Evry
 France
olivier.michel.fr@gmail.com
17. **Sonja SMETS**
 Research Group in Philosophy of In-
 formation, GPI
 University of Hertfordshire
 College Lane
 Hatfield AL10 9AB
 United Kingdom
sonsmets@vub.ac.be

Statistical information on participants

Before we present the usual statistical information in the table below, note that it was extremely important for the success of this event to have the required variation of expertise in the participant pool. The diagram shown here was presented on the workshop website early on, and summarises the (sometimes overlapping) skills of all participants. Not all their skills of course, just the four important ones for the workshop. The rough criterion for being in a set is to have written at least one technical paper on the subject. Note that everybody has had experience with using and developing programming paradigms, a definite asset to the success of the workshop. Note that the diagram does not include the ESF representative.



Total number of participants		17		Country of residence	
		Belgium		2	
		Denmark		1	
Sex		Finland		1	
Male	12	France		4	
Female	4	Germany		1	
Age		Italy		1	
< 30	1	Sweden		1	
30-40	8	UK - England		3	
40-50	2	UK - Scotland		2	
50-60	4	USA		1	