Background Reading for ACCS Winter School

Topic 3: How complex is the sharing of non-point source pollution costs?

Presenter: Rodney Beard
Assistants: Stuart McDonald
Liam Wagner

Basic problem: Pollutants such as nutrients and pesticides from agricultural run-off flow down a river system and pollute the coastal marine environment. The impact of this on coastal marine economic activities needs to be quantified in terms of the **opportunity costs** to stakeholders who participate in economic activity that depends on the quality of the coastal marine ecosystem. Such stakeholders include, fishermen both commercial and recreational, tourist operators and tourists and ancillary service industries associated with these activities. Costs to stake holders can be determined either by direct measurement after the damage is done, or by computational modelling of the linkages between coastal agriculture and the marine ecosystem. Such models involve for the most part techniques drawing on non-linear dynamics and optimal control theory and are related to questions of ecosystem resilience.

This problem is a non-point solution problem in which it is impossible to identify precisely the source of the pollutant. However, it is possible to share the costs amongst those who are collectively responsible for the damage. Once the opportunity costs to stakeholders have been determined they need to be allocated amongst the perpetrators of the **externality**. This needs to be done in as fair a manner as possible so that principles of distributive justice are adhered to. In the case of pollution from coastal agriculture the costs of the pollution need to be shared amongst the farming community along the river. The community of farmers in a river basin may be viewed as a network of agents. This network is a combination of the physical river network and the **cadastral** land-use network of the farmers. River basins may be viewed from a complex systems perspective as fractal river networks, characterized by properties such scale-freeness, power law relationships, etc. The socio-economic network results from superimposing the cadastral properties of a geographic region on the river network.

How do we address the problem of cost-sharing amongst agents located on such networks and how complex is this problem? The literature on cost-sharing in economics primarily employs cooperative or coalitional form game theory as an analytical tool. Game theory is a branch of mathematics that is used to study the interaction of multiple agents engaged in economic, social, biological, engineering and physical activity. As a theory of multi-agent behaviour it is of clear relevance to complex systems science.

In the previous session you looked at games in which cooperation between players (agents) is not binding and communication during play is not allowed. We will now introduce some basic terminology of the case where communication is allowed and cooperative agreements are binding.

Consider a set of *n* agents $N = \{1,...,n\}$ this is called the **player set**. The potential set of N is denoted by 2^N is the set of all possible coalitions S of N. We can now define what we mean by a **game**.

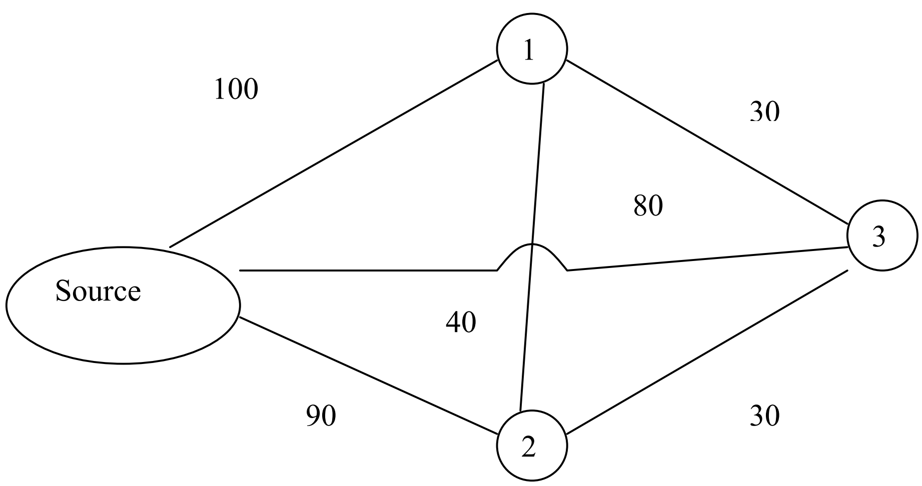
Definition 1 (Game) A cooperative n-person game in coalitional form is an ordered pair (N, v), where $N = \{1,...,n\}$ and $v : 2^N \rightarrow \Re$ is a map assigning to each coalition $S \in 2^N$ a real number *v*, such that $v(0) = 0$ (Tijs 2003: 60-61).

The function v is called the characteristic function, the worth or the value of a coalition. In cost-sharing games this value is a cost to be shared amongst the members of the coalition.

Here is an example adapted from Tijs (2003: 61).

E.1 Three cooperating communities.

Communities 1,2, and 3 want to be connected to a nearby water supply (dam or river). The possible pipelines are given by the following network:



The weights on the graph are costs of pipeline construction.

Cooperative games are of two broad types transferable utility (TU)-games and non-transferable utility (NTU)-Games. TU-Games assume that utility of agents is quasi-linear, that is made up of a linear and a non-linear component. The linear component involves a monetary transfer between agents.

Cost Allocation Games

Let's look at some examples of cost allocation games**.**

E.2. Municipal cost-sharing problems

A group of N towns are considering the cost of constructing a shared water treatment facility. Each town has a minimum water requirement that can be met either by themselves or through a water sharing agreement. $c(S)$ is the minimum cost of supplying members of a coalition of towns by the most efficient means.

E.3. Airport Games

Consider an airport with one runway. There are m different types of aircraft. c_k is the cost of constructing a runway for an aircraft of a given type *k*. N_k is the set of aircraft landings of type k in a given time period. The players of the game are landings of aircraft. The cost function of the airport game is

$$c(S) = \max\{c_k \mid S \cap N_k \neq 0\} \text{and } c(\phi) = 0$$

the airport game was origninally formulated for the calculation of landing fees for Birmingham airport, UK in the 1970's. the original problem involved 11 different aircraft types and 13,572 players (aircraft movements). Landing fees were calculated based on a concept known as the **nucleolus**. The solution algorithm employs linear programming. This problem demonstrates the applicability of game theory to solving large real world multi-agent cost allocation problems.

E.4. Minimal Cost Spanning Tree Games

A group of n customers spread through a geographical region need to connect to a supplier of some resource. Define $N_* = N \cup \{0\}$. Consider a complete undirected graph with a set of nodes given by N_*. Customers can connect to the supplier directly or indirectly. Consider the complete undirected graph whose node set is $N_* = N \cup \{0\}$. The cost of adding an edge $e_{\{i,j\}}$ connecting nodes i,j in N_* is given by $c_{\{i,j\}}$. Given $S \subseteq N$ a minimum cost spanning tree is given $\Gamma_S = (S \cup \{0\}, E_S)$ is a tree with node set $S \cup \{0\}$ and a set of edges E_S that connects members to the common supplier such that the set of total costs is minimal.

The total cost function is given by $c(S) = \sum_{e_{\{i,j\}} \in E_S} c_{ij}$ for all $S \subseteq N(c(0) = 0)$.

Interpretation: consider the resource to be the service the ecosystem supplies as a sink for pollutants. The supplier node is the rivermouth or marine ecosystem. The edges are pollution pathways into the marine ecosystem. The costs associated with each edge are costs of eradicating pollution along a particular pathway such as the branch of a river. Can we allocate the cost of polluting minimally amongst players located along the river?

Note minimal cost spanning tree games generalize airport games.

E.5. Fixed cost spanning forest games

Note: fixed cost spanning forest games generalize minimum cost spanning tree games.

Efficiency

The core. The core of a game (N, v) is the set:

$$C(v) := \left\{ x \in I(v) \mid \sum_{i \in S} x_i \geq v(S) \quad \text{for all } S \in 2^N \setminus \{\phi\} \right\}$$

where $I(v)$ is the set of imputations. An element of this set is a payoff distribution of the earnings of the grand coalition. This means that no coalition S has an incentive to split off because the sum of payoffs to members of the coalition under a proposed imputation (payoff distribution) is at least as much as what the coalition members could get if they were to split off.

Fair sharing

The Shapley value is given by:

$$x_i = \sum_{s=0}^{n-1} \sum_{S \in A_i(s)} \frac{s!(n-s-1)!}{n!} \{C(S \cup \{i\}) - C\{S\}\}$$

The Shapley value is the average marginal contribution of a player to a collation over all possible orderings of the coalition. The Shapley value has an axiomatic basis the key axioms being:

1. equal treatment of equals (fairness principle)
2. dummy axiom (agents who contribute nothing pay nothing)

3. additivity (decentralization property – essential for distributed systems)

Myerson value is a version of the Shapley value defined on graphs, i.e. agents are related to each other by a graph or network.

Here is a computational example:

Consider the problem of 3 agents sharing joint costs
Example: 3 residents of a sharehouse sharing rent
Who should pay how much rent? The residents are called A,B,D

- Each of them pays 200 rent per week to rent by themselves
- If they share a house they pay 350 rent between them
- If A and B or A and D share they can rent for 200
- If B and D rent they pay 350

Table 1

order	A	B	D
ABD	200	0	150
ADB	200	150	0
DAB	0	150	200
DBA	0	150	200
BDA	0	200	150
BAD	0	200	150
Average	66.67	141.67	141.67

The Myerson value – an analogue of the Shapley value for graphs will be covered in the lab session.

Complexity

How does this relate to complexity? The combinatorial nature of games on networks makes these games difficult to solve. Solutions to cost-sharing problems vary in how difficult they are to compute. If the solutions are too complex then only simplified versions of the problem can be solved or they will have have to be computed using meta-heuristic methods, such as those used in evolutionary computation, e.g. genetic algorithms, directed search methods, simulated annealing, Tabu search and other techniques.

The complexity of solution concepts for cost-sharing problems is not known in all cases.

What do we mean by complexity?

There are three broad notions of complexity that can be identified and that can also be related to each other. The first is **algorithmic complexity** and the second **informational complexity (Kolmogorov complexity)** and thirdly **information based complexity**. Schirmer (1995:2) defines the "inherent computational tractability of problems" as the central subject matter of complexity theory. This would hold true of all three approaches to complexity.

Algorithmic complexity is concerned with the computational effort required to solve a particular problem by implementing a given algorithm. The basic idea is that the time complexity of an algorithm is some function of the input size. Algorithms for which this function is a polynomial with respect to input size are said to be tractable. The class of tractable problems is referred to as **P**. The class of problems that can be solved by a nondeterministic algorithm, e.g. some form of random search, is known **NP** (non-deterministic polynomial). Problems are either tractable or intractable. However, it is often not known whether a problem is tractable or intractable. Examples in this class include NP-complete problems. A problem is NP-complete if every other problem in NP can be transormed via some polynomial to that problem. A problem that can be polynomially reduced to an NP-complete problem by this means is referred to as NP hard. Problems in this class are amenable to analysis via meta-heuristic methods, such as random search, evolutionary programming and other similar non-deterministic search and optimization techniques. Hence the popularity of these methods in complex systems science. Many other complexity classes are known that will not be discussed here.

Informational complexity falls into two categories: Kolmogorov complexity which analyzes the informational cost of computation using a Turing machine model of information processing and information based complexity (IBC) which uses a real number model of computation. Kolmogorov complexity and algorithmic complexity are appropriate for discrete problems and information based complexity for continuous problems.

Algorithmic complexity can typically be reduced to Kolmogorov complexity via the Turing machine model.

Kolmogorov or Kolmogorov-Solomonoff complexity is given by:

$$C_f(x) = \min \{l(p) : f(p) = n(x)\}$$

Where $f(p)$ is a partial function defined over integers, $l(p)$ the length of a given program and $n(x)$ the size an object x, where x is a finite string or message. This definition of complexity defines the complexity of a message as the shortest program capable of processing a message of a given size. It is usual to think of *f(p)* as a Turing machine. The Turing machine model works for integer valued problems and problems in which the object is measurable so that *n(x)* has a straightforward interpretation. For continuous problems information based complexity is appropriate rather than the Turing machine model.

Information based complexity is that branch of complexity theory which studies problems in which information is either partial, contaminated or priced and involves determining the minimal computational resources necessary for solving continuous problems. Information based complexity does not employ the Turing machine model as as the basis for computation. Instead, information based complexity employs the so-called real number model of computation (think here of floating point operations in scientific computing). This approach assumes:

1. arithmetic operations and comparisons on real numbers can be performed at unit cost
2. information operations can be performed at cost c>>1.

In information based complexity theory the tractability of a problem is defined similarly to AC and K-SC theory. A problem is said to be tractable with respect to a particular parameter if the problem complexity grows at most polynomially with respect to the parameter (Traub and Werschulz, 1998: 27).

How do we define complexity in this sense.

Consider the solution *S* to a problem. The solution can be considered a mapping from one normed linear space *F* to another *G*. Consider an element *f* of the problem space *F*.

Following Traub and Werschulz we now define an information operator

$$N(f) = [L_1(f), \dots, L_n(f)]$$

of permissible information functionals (L_i). An algorithm is now a mapping ϕ from the set of information functionals to the solution space G. for any problem f we can now compute an approximation $U(f) = \phi(N(f))$ to a solution *S(f)*. To do this we define the error of a given approximation by:

$$e(U, f) = \| S(f) - U(f) \|$$

Note this is why we needed to assume F,G to be normed linear spaces.

The cost of an approximate solution to a problem is given by:

$$\cos t(U, f) = \cos t(N, f) + \cos t(\phi, N(f))$$

where cost is the sum of the cost of obtaining and cost of combining information.

The computational complexity of an approximation is given by:

$$comp(\varepsilon) = \inf\{\cos t(U) : U \text{ such that } e(U) \leq \varepsilon\}$$

The complexity is thus the smallest cost of an approximation such that the error remains within a given tolerance.

Costs and approximate solutions may be defined differently for different algorithms.

We can now define what we mean by complex systems.

Complex systems are systems that generate problems that are computationally intractable.

Whereby tractability may be defined based on complexity in any of the above senses.

In what follows most of the complexity results reported fall into the algorithmic complexity category. Although some problems can be examined from the information based complexity perspective. All other definitions of complexity will be reduceable to one or another of the above definitions.

A number of complexity results are known for solution concepts for cooperative games. An interesting survey is given by Bilbao et al. in what follows I will summarise some of these results (see tables 1 and 2).

Table 2

Game	Concept	Result
wGG	Core membership	P
MBG,MFG	“	?
wGG	Core empty	P

MFG	Core empty	P

Table 3

Game	Concept	Result
MSTG	nucleolus	NP-hard
MFG	subadditivity	O(size(v))
MBG	submodularity	Oracle-polynomial time
MFG	submodularity	O(size(v))

Fractal Networks

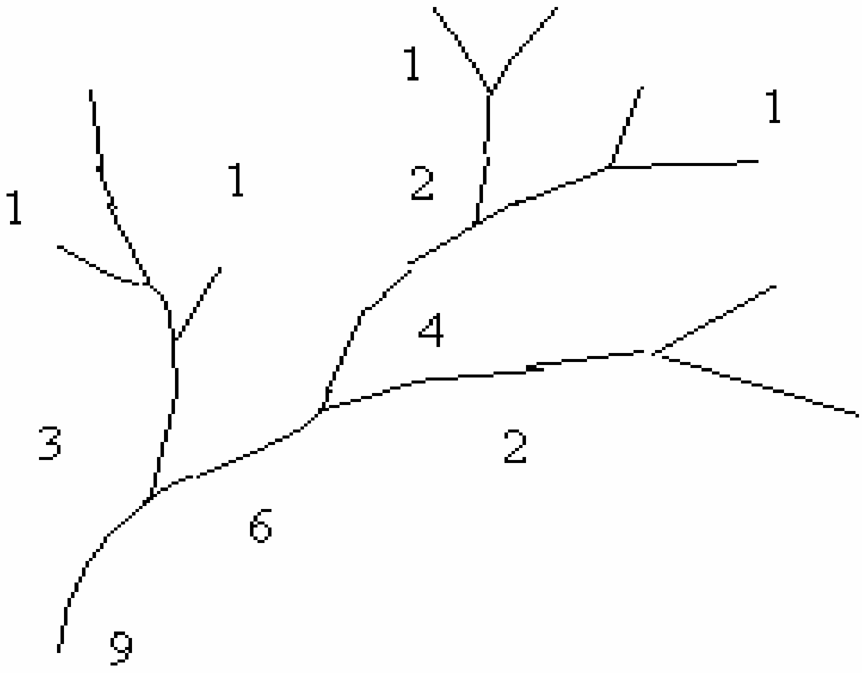
Why are power laws, scale-free properties and self-organized criticality thought to be properties of complex systems? Power law and scale free properties are related to fractal properties of systems.

The example we will look at here is river networks. Complex systems properties of river networks are well established. Horton (1945) developed a scheme for indexing streams according to their hierarchy. Strahler (1957) modified Horton's ordering scheme.

The Horton-Strahler ordering proceeds as follows:

1. Each source is assigned the order i=1
2. the order of successor streams is determined by the orders of the predecessor streams. Given predecessor streams $i_1 \neq i_2$ then $i = \max\{i_1, i_2\}$ otherwise $i = i_1 + 1$.

A number of properties of river networks can be computed from the Horton-Strahler ordering. Consider the number of streams of a given order N_i. then the bifurcation ratio can be estimated by the bifurcation ratio: $B_i = N_i / N_{i+1}$.

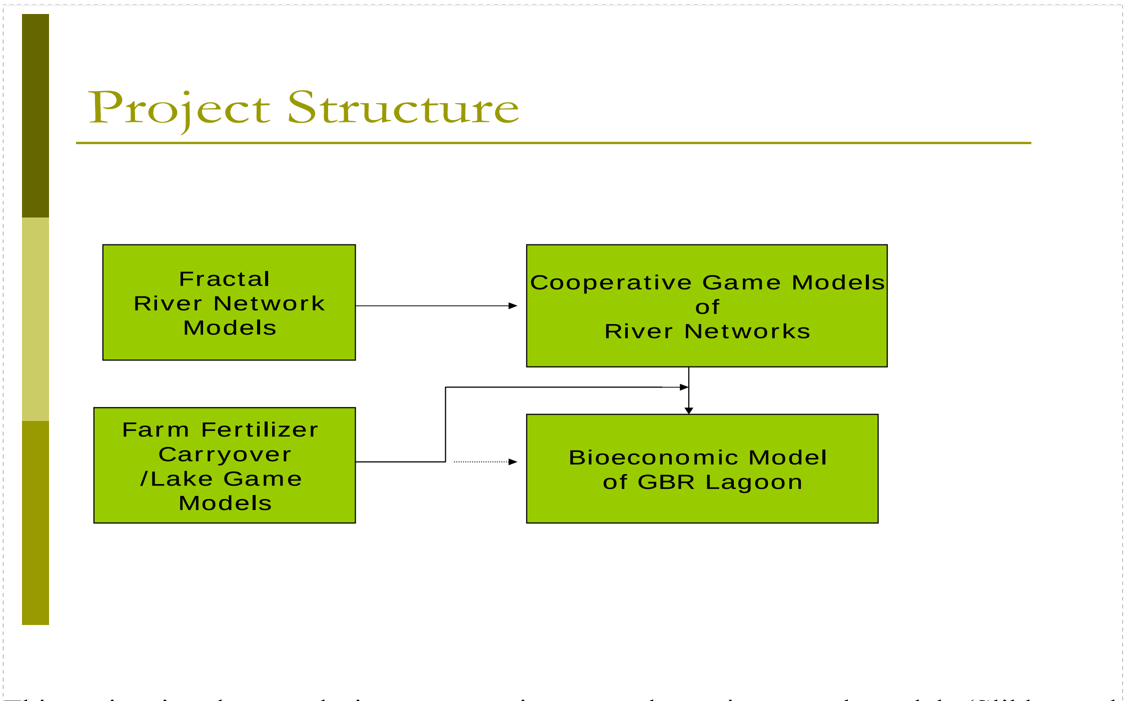


Exercise : Calculate the bifurcation ratio for this network!

Horton's law of stream lengths is given by $R_L = {L_i}/{L_{i-1}}$ and Hortons law of areas is given by: $R_A = {A_i}/{A_{i-1}}$ where A refers to the are of a basin of particular stream order and L to the length of streams of a particular stream order. The fractal dimension of a river network can then be calculated by $D = 2\frac{\ln R_L}{\ln R_A}$. Hack's law relates area and length of river basins by a power law relationship: $L = CA^{\alpha}$. Hack's law assumes rivers are not elongated. For elongated rivers it can be replaced by the idea of Tokunaga cyclicity.

A number of toolkits exist for computing key parameters of river networks for example Tardem (http://hydrology.neng.usu.edu/taudem/) and CatchmentSim (http://www.toolkit.net.au/catchsim/).

Eutrophication of the Barrier Reef Project



This project involves exploring co-operative game theoretic network models (Slikker and van den Nouweland, 2001) and combining them with models of river systems (Rodriguez-Iturbe and Rinaldo, 1997) in order develop socio-economic models of riverine communities**.** The plan is then to combine these with bioeconomic models of the Barrier reef lagoon. The bioeconomic models will be used to obtain cost estimates of the impact of pollution damage on the reef lagoon. The project structure is illustrated above.

Evolution and Dynamics

These games are often seen as largely static in nature, however there is an emerging literature on dynamic coalitional form games. I will give a brief overview of this literature. Which studies how characteristic function evolve over time. In applying this to cost-sharing this dynamic coalition form goes show how cost-sharing rules evolve through time and need to adapt to changes in the state of a the system.

Some of the relevant literature includes:

Dementieva, M. (2004) Regularization in Multistage Cooperative Games, PhD Thesis, Jvaskala Studies in Computing No. 42, University of Jyvaskala.

Filar, J.A. and L.A. Petrosjan (2000) Dynamic Cooperative Games, International Game Theory Review Vol. 2, No. 1, pp. 47-65.

Petrosjan, L.A. and N.A. Zenkevich (1996) Game Theory, World Scientific,

Singapore.
Petrosjan, L.A. and G. Zaccour (2003) Time Consistent Shapley Value Allocation of Pollution Cost Reduction, Journal of Economic Dynamics and Control, Vol. 27, pp. 381-398.
Petrosjan, L.A. (2005) Cooperative Differential Games, in: A. Nowak, K. Szajowski (Eds.) Advances in Dynamic Games: Applications to Economics, Fianance, Optimization and Stochastic Control, Annals of Dynamic Games 7, Birkh¨auser.
Tarashnina, S. (2002) Time-Consistent Solution of a Cooperative Group Pursuit Game, International Game Theory Review Vol. 4, No. 3 pp. 301-317.
Zakharov, V and M. Dementieva (2004) Multistage Cooperative Games and the Problem of Time Consistency, International Game Theory Review, Vol. 6, No. 1, pp. 156-170.

In our own work we have extended the water resource allocation game of Ambec and Sprumont (2002) to a multistage setting using these techniques.

Readings

E. Algaba et al. On The Complexity Of Computing The Myerson Value By Dividends (http://citeseer.ist.psu.edu/518438.html)

S. Ambec and Y. Sprumont (2002) Sharing a River, *Journal of Economic Theory* 107: 453-462.

J. M. Bilbao, J.R. Fernandez and J.J. Lopez Complexity in cooperative game theory, mimeo.

G. Chartrand and O.R. Oellermann *Applied and Algorithmic Graph Theory*, McGraw-Hill New York 1993.

U. Faigle et al. On the Complexity of Testing Membership of the Core in Min-Cost Spanning Tree Games, (http://citeseer.ist.psu.edu/faigle94complexity.html)

M.R. Garey and D.S. Johnson *Computers and Intractability: A Guide to the Theory of NP-Completeness*, San Francisco: W.H. Freeman 1979.

L.A. Hemaspaandra and M. Ogihara *The Complexity Theory Companion*, Berlin Springer Verlag, 2002.

T. Ichiishi *Game Theory for Economic Analysis*, New York: Academic Press, 1983.

M. Li and P. Vitanyi *An Introduction to Kolmogorov Complexity and its Applications*, Berlin: Springer-Verlag 1997.

R. B. Myerson Graphs and cooperation in games, *Mathematics of Operations Research ,* Vol 2 No. 3 August 1977.

G. Owen, *Game Theory*, San Diego: Academic Press, 1995.

G. Owen, *Discrete Mathematics and Game Theory*, Boston, MA: Kluwer, 1999.

C. Papadimitriou *Computational Complexity*, Reading, MA: Addison-Wesley 1994.

B.Peleg and P. Sudholter *Introduction to the theory of cooperative games*, Kluwer, 2003.

Ignacio Rodriguez-Iturbe, Andrea Rinaldo *Fractal river basins : chance and self-organization* Publisher Cambridge ; New York : Cambridge University Press, 1997

A. Schirmer *A Guide to Complexity in Operations Research* (http://halfrunt.bwl.uni-kiel.de/bwlinstitute/Prod/mab/schirmer/research/pdf/wp381.pdf)

W.W. Sharkey Network Models in Economics, in: M.O. Ball et al. (Eds.) *Handbooks in OR&MS*, Vol. 8. Elsevier 1995.

S. Tijs *Introduction to Game Theory*, Hindustan book Agency 2003.

J.F. Traub and A.G. Werschulz *Complexity and information*, Cambridge University Press 1998.

S. Watanabe *Kolmogorov complexity and computational complexity*, Berlin: Springer-Verlag 1992.

Key concepts

Cadastral

derives from the French *cadastre* and originally from the Greek *Katastikhon* and defines the extent, value and ownership of land for taxation purposes (http://www.nla.gov.au/map/cadastral.html)

Complete graph

Externality

Opportunity costs