Complex Buildings and Cellular Automata—A Cellular Automaton Model for the Centquatre-Paris

Roberto D’Autilia 1,* and Janet Hetman 2

1 Dipartimento di Matematica e Fisica, Roma Tre University, Largo S. Leonardo Murialdo 1, 00146 Roma, Italy
2 Dipartimento di Architettura, Roma Tre University, via Madonna dei Monti 40, 00184 Roma, Italy;
janet.hetman@uniroma3.it
* Correspondence: roberto.dautilia@uniroma3.it

Received: 21 May 2018; Accepted: 11 June 2018; Published: 14 June 2018

Abstract: We explored the relational dynamic elements of complex buildings, a type of architecture designed to incubate uses, located in urban areas with high housing density. The uses of complex buildings concern different elements, including the network of agents using or managing them, the environment, and the activities and functions that take place occasionally, temporarily, or permanently. Data was gathered through ethnographic research lasting 6 months, and a chronotopian approach was used to describe time and space. We analysed and discussed the interaction of the elements of complex buildings through a cellular automaton model, a computational method that simulates the growth of complex systems. It was used here to generate patterns that suggest configurations of uses that can optimize management and therefore increase economic and social capital. The cellular automaton representation was used to develop an abstraction of the Centquatre, a public cultural center in Paris. This center is a good example of a complex building, being based on a public–private partnership and having an architectural configuration designed to host a wide range of art, social, and productive activities. The building includes a large central space used as an urban public area open to different types of people. The relevance of the case study lies in its capacity to produce economic value by combining different uses, and also by welcoming different people to the public space. We found that the multistate cellular automata representation allows the Centquatre behavior to be modeled by means of combinatorial and statistical methods. The correlations between the automaton behavior and the number of users can be identified using machine learning techniques related to random forests. We argue that this approach makes it possible to improve the planning of complex buildings.

Keywords: architecture; complex buildings; cellular automaton; uses incubators; public space

1. Introduction

In a memorable essay addressed to biologists, John von Neumann describes the structure of an automaton that reproduces itself, even in more complex or evolved forms, illustrating the logic that serves as a deciding principle for living beings and computers [1]. This idea has produced many results, showing how a simple local rule can produce complex global behavior. It is the main idea behind cellular automaton theory.

These concepts have been widely applied in different fields, from statistical mechanics to traffic analysis and models of artificial intelligence. The paradigm has produced a large range of results and methods that are now used to model and understand complex collective phenomena. The increasing speed of computers has also made it possible to simulate the dynamics of structures composed of...
many interacting subparts efficiently and realistically. Some formal methods of analysis, often derived from other disciplines, have been widely used in urban planning \[2,3\]. On a smaller (e.g., architectural) scale, these methods seem more difficult to incorporate \[4\]. The large numbers of urban scale lead naturally to probabilistic analysis, while large numbers are generally difficult to find in the behavior of a single building. However, probability analysis has been used to generate architectural forms for single buildings \[5,6\]. Many functional programs in architecture involve distributing spaces and activities so as to optimize surfaces, resources, and efficiency of use, even in complex buildings \[7\].

The increase in complexity at architectural scale is well-expressed by recent history. Buildings in dense urban environments have been designed to allocate and accommodate multiple uses in a space-time program. Architecture has recently dealt with this problem in connection with mixed socio-cultural buildings. The well-known design experiences of the Fun Palace by Cedric Price, the Parc de la Villette (which embodies the Event-City) by Bernard Tschumi, and the Centre Pompidou by Renzo Piano and Richard Rogers \[8–10\] indicate a family of buildings known today as “complex buildings” \[11,12\]. While cities have witnessed a proliferation of practices \[13\], buildings have accommodated a multiplication of uses. These buildings become incubators of uses, in which the complexity involves a multitude of configurations with different uses, actors, users, equipment, etc. The consequence is an increase in the number of variables and the development of new forms of management—buildings are transformed into intelligent systems that meet the dwelling needs of the contemporary city.

From this point of view, buildings must be programmed to ensure good cohabitation of their different uses. Management must be able to arrange (program) the uses in spatio-temporal configurations in which the main goal is to preserve the mutual compatibility of the parts \[14\]. However, the many possible configurations of a multifunctional or complex space suggest the use of statistical methods for its planning. It is therefore necessary to exploit some machine learning methods to establish the correlations between configurations and responses. Representing the complex building as a cellular automaton, the comparison between these methods shows that the best algorithm to correlate configurations with uses is the random forest \[15\]. The main results of our analysis are twofold: we show that a multifunctional building that gives rise to a gigantic number of possible configurations can be studied by statistical-combinatorial techniques, and also that the planning of such spaces can be optimised by means of the random forest method \[15\].

2. The Centquatre-Paris

The Centquatre (établissement artistique et culturel) located in the 19th arrondissement of Paris (Figure 1) is the French case study chosen \[16,17\]. As a part of the French political program for the decentralisation of culture, the City of Paris transformed the nineteenth-century public undertaker’s building \[18\] into a public cultural center in 2008. Formally, the Centquatre is an EPCC (établissement public de cooperation culturel) which enables it to build partnerships with other public or private entities for cultural, artistic, and social purposes \[19\].

Its main activities include free social activities for those living in the neighbourhood (adults and children), artistic production, shows, exhibitions, and commercial and productive activities (shops and restaurants; a start-up incubator; an architectural, urban planning, and cultural engineering consultancy service). These activities address and attract very different persons. The Centquatre’s population is therefore a mix, including residents of the district, people from the city center or suburbs, artists and users of artistic activities, designers and the like, hosted in the factory. The institute receives a variety of people that “reproduces the world”, in the words of one director, Jean Bourbon.

In order to better understand the complex phenomena of the Centquatre, we start by describing the building as an incubator of uses, showing how the different elements relate to each other.

We propose two descriptions given by the main managers: the first Christophe Girard, president of the board of directors and promoter of the project as councillor for culture at the time of the restoration, and the second José-Manuel Gonçalvès, sociologist and artistic director of the Centquatre.
Christophe Girard defines the institute as a collaborative artistic platform in which art is made accessible thanks to a popular, yet contemporary and committed, program. For him, the Centquatre is an “aesthetic refuge where art, society and innovation meet”, and where there are “multidisciplinary arts, public mixité, diversity of activities between art and innovation, commerce and public interest. The links between these activities allow for almost infinite combinations” [20].

José-Manuel Gonçalvès emphasises that the institute’s goal is to be a place of “experimentation, rather than a model, where the primary interest is aimed at doing together” [20].

“Experimentation on the links” between the different activities guarantees that the Centquatre is conceived, managed, and used as an “open place”, a space available for individual and collective uses, and in which subjects, actors, and activities interact dynamically. The institute is committed to simultaneous actions at different levels in an integrated way:

- Social: LeCinq, La Maison de Petits, and Public Space
- Artistic: shows, concerts, exhibitions, activities
- Mediation with the public
- Production and innovation: performing arts, encouraging artistic activity by the non-expert public, start-up incubator, consultancy
- Management: planning, relationships with sponsors and customers
- Commerce: restaurants, retailing, rental space for events.

The set of actors, activities, and spaces indicates the vast range of interaction possibilities the Centquatre offers its users and the city. The program attracts people who shape the platform that determines the components.

The institute and its spaces are attractive by virtue of the program, which is the determining component of the configurations. In fact, the activities are organised on the basis of the envelope and the furnishing necessary for each event. Besides its own activities, the Centquatre also hosts private events, offering rental premises (fashion showrooms, conventions, catering, etc.). This availability is profitable and simultaneously attracts very different people and uses.
Since the activities do not have a foreseeable recurrence, the program is defined week by week. High uncertainty means that the management has to set up different configurations, organising activities according to their compatibility. Thus, the Centquatre can be considered a complex dynamic machine in process, an “open place”: places that reinvent themselves with each new initiative [20].

This plurality of uses is possible because of architecture designed especially for this purpose (Figure 2). In fact, the restoration was the result of a French procedure known as Marché de définition [21], in which the functional program and the architectural requirements were defined simultaneously. Usually applied to large urban areas, this procedure made it possible to configure the space so as to increase the possibilities of the building thanks to the large central space. In this “yard”, used as a public area, different activities can take place at the same time. Each attracts a particular public, so that all the people participating in given activities are concentrated here (Figure 3). Like an urban public space, the yard becomes a place where relationships can be formed and cultivated. The resource that the Centquatre attracts is participation of the public in events taking place in this public space. The aim of the management is to multiply the physical presence of the people in the public space to generate more value.

Figure 2. Plans and longitudinal section of the Paris Centquatre building.
We discuss the value added by the physical presence of people in the space, the way in which the public space is the real resource of the center. This architectural space guarantees simultaneous integrated development of the various activities that take place in the 39,000 square meters of the building. The public space of the central yard plays a paramount role in achieving this.

3. The Data

This case study was chosen because it is highly innovative, involving up to 12 categories of activities and attracting many visitors (about 5000 per day). These figures are further supported by the increasing budget of the Centquatre in the last two years (+25% on a budget of €16 million per year). In addition to management skills, the architectural configuration helps to obtain these numbers, basically because the main central space is configured and used as a public area.

Our study takes a mixed approach [22,23], in which quantitative and qualitative tools are used in a productive, integrated way [24]. The cellular automaton representation presented here would not have been possible without ethnographic knowledge [25]. The collection of qualitative data was based on 6 months of observations, unstructured interviews, and a field survey of persons frequenting the center. In order to make qualitative data suitable for use with the cellular automaton model, daily programming tabs and reports were classified in quantitative categories.

Ethnographic analysis provided empirical knowledge that facilitated comprehension of the phenomena and guided interpretation of the field data (i.e., the daily planning schedules; Figure 4). A chronocarta, that is, an experimental dynamic map of a complex building in its space-time component [26,27], was constructed with this qualitative data (Figure 5).

In addition, the center collects data on daily inflows of the public to the activities of the program (Figure 6). These two data sources were analysed as cellular automata configurations of an unknown process.
Figure 4. Planning sheet of Centquatre for 18 March 2017 ©CENTQUATRE-PARIS.

Figure 5. Chronocarta for 27 February 2017, 5 to 6 p.m. The building plan is a symbolic representation of its usage.
The collected data were provided by the Centquatre board during the six months of ethnographic study (October 2016 to March 2017). A fundamental step in a mixed-method approach is to select and edit data from different origins to make it homogeneous and valid over the study period. The configurations analysed report a total of 26,694 people present during the 58 configurations. The average was 460 people per configuration, with a minimum of 0 people during the days of any activities, and a peak of 3755 people because the cultural center hosted a large and public event. The data came from three sources:

- The Chinoiserie (urbanfurniture): about 75
- Nicolas Clausse (free exhibition): about 140
- Il était plusier fois (Theatre Show): 61
- Circulation(s) (photographic expo fee and free): 226 (ticket numbers)
  - Atelier 0 (P): 142
  - Atelier 2 (P): 197
  - Ateliers 1, 3, 5 (P): 224
  - Halle Aubervilliers (Free): 337
  - Little Circulation (Free): 205

- Les Illuminations
  - Thursday: 66
  - Friday, 10 am: 117
  - Friday, 2:30 pm: 54

- Toutes les femmes sont des Aliens + Les oiseaux reviennent (Theatre Play): 56

Figure 6. Daily report for 3 February 2017 ©CENTQUATRE-PARIS.

The collected data were provided by the Centquatre board during the six months of ethnographic study (October 2016 to March 2017). A fundamental step in a mixed-method approach is to select and edit data from different origins to make it homogeneous and valid over the study period. The configurations analysed report a total of 26,694 people present during the 58 configurations. The average was 460 people per configuration, with a minimum of 0 people during the days of any activities, and a peak of 3755 people because the cultural center hosted a large and public event. The data came from three sources:

- The online program which mainly included “Que faire au 104”, activities and events open to the public, often free of charge.
- Centquatre internal planning data including spatialized events (assembly and development).
- Centquatre daily reports describing the status of the activities and the number people attending.

To further optimize the analysis, the selection of data and calculation only concerned configurations of February and March 2017, a period characterised by data with good continuity and homogeneity. Although we lacked the daily reports for one week, we estimated and reconstructed the missing data to enable quantitative analysis.

The data was summarised and coded during the transcription process, and numerically described the configurations of each active hour of a single day. A total of 53 days were analysed, with an average of 11 h per day. This resulted in 583 real configurations, from which the most significant configurations for each day were extracted and further analysed. The calculation therefore showed 10% of the data gathered.

The program data was differentiated into sheets for each day, and as shown in Figure 4, these sheets contained information on the activities allocated by temporal extension and spatial envelope. The activities were differentiated by color bands labeled with the name of the activity and further details on the specific location. The idea of the chronocarta was to make the planning scheme more comprehensible: it simplified reading of the elements and brought out the dynamism of the configurations in the flow of space-time.
The daily reports included:

- Turnout to shows. This number refers to the specific activity, and the total refers to the day and is specific to the space section.
- Number of participants in free activities open to the public. This figure sometimes refers to the specific activity and sometimes to all activities. Also in this case, the total refers to the day and is specific to the space section.
- Number of people in the public space. The figure refers to all spontaneous activities occurring in the public space. In this case, the reference figure is counted over the week with a distinction between weekdays and weekends. It does not specify a spatial section but the area of public space available, which is not always the same.

The study period captured the range of the activities taking place in the center. On Mondays the center is closed to the public, although activities may still occur. This is why we did not collect daily data on the use of the structure, and why Monday configurations mostly have a negative value. In parallel to the data editing, a simplified drawing of the building was made.

4. Description of the Methods

The simulations were carried out using the analytical tool of Mathematica 11.2 software. The starting model was the cellular automaton. The building was seen as a set of cells, each of which can be in the state corresponding to its purpose (Figure 7).

<table>
<thead>
<tr>
<th>Service / Ticket</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proximity services</td>
<td>2</td>
</tr>
<tr>
<td>Commercial activities</td>
<td>3</td>
</tr>
<tr>
<td>Exhibitions</td>
<td>4</td>
</tr>
<tr>
<td>Shows</td>
<td>5</td>
</tr>
<tr>
<td>Groupes</td>
<td>6</td>
</tr>
<tr>
<td>Inaccessible activities</td>
<td>7</td>
</tr>
<tr>
<td>Private events</td>
<td>8</td>
</tr>
<tr>
<td>Partnership</td>
<td>9</td>
</tr>
<tr>
<td>Internal visits / meetings / activities</td>
<td>10</td>
</tr>
<tr>
<td>Public Space / Spontaneous practices</td>
<td>11</td>
</tr>
<tr>
<td>Toilet / Connective facilities</td>
<td>12</td>
</tr>
</tbody>
</table>

**Figure 7.** The table of uses for Centquatre together with the numerical and color code used for the analysis.

The center’s spaces are divided into 67 cells. If each is in one of the 12 possible states, the possible configurations of purpose of the center are $12^{67}$, a gigantic number of the order $10^{72}$. Of course, not all areas can have all of the functions, but even if the number of targeted configurations is reduced, it is large enough to support statistical analysis.
It should be added that a rule for the cellular automaton states that the change of state of a cell is influenced by nearby cells. In our case, modification of the state of an area depends on the state of all other areas. Some of these dependencies are strong, some weak, some conflicting, but for the moment, it is not known how this complex network of relationships is structured. It is known that a reward is associated with each configuration of real purpose: a score, such as the number of visitors or an economic income. Based on the available information (i.e., actual configurations and corresponding rewards), we set out to estimate the network of relationships that binds individual areas. From this estimate, we calculated the possible reward of a configuration not yet tested.

As a reward for a configuration, we chose the number of people visiting the building in that condition. This number can of course depend on many other variables, such as weather, advertising, or a specific attraction in a space in the building. However, we were constructing a statistical regression model and therefore assumed that for large numbers these variables were mixed somewhat randomly, partly determining probabilities.

We divided the area of the center into 352 squares by superimposing a square lattice of dimension 3211, and we associated a color and a numerical code corresponding to its function with each square, according to the numbers–colors codes of Figure 7. In this way, each configuration of purpose of the Centquatre can be represented graphically. The Figure 8 shows a sample of 16 real configurations with the corresponding number of visitors.

![Figure 8. Configurations with corresponding numbers of visitors.](image)

To estimate how the different purposes are correlated with each other and with the number of visitors, we exploited four machine-learning models: linear regression, nearest neighbours, neural network, and random forest. The aim was to determine the most reliable model. For descriptions of the functioning of these models, readers are referred to the abundant literature on the subject [28,29].

We trained the four machine-learning models to associate 50 real Centquatre configurations with the corresponding number of visitors, and then used predictors to predict the number of visitors for the six real configurations of Figure 9.

![Figure 9. The test configurations with the corresponding number of visitors.](image)
The number of visitors expected for these configurations, according to the four models, are shown according to Figure 10.

![Figure 10. The prediction table for the five configurations used to test the methods.](image)

<table>
<thead>
<tr>
<th>Predictor Information</th>
<th>Discrete Predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real</td>
<td>286</td>
</tr>
<tr>
<td>Linear Regression</td>
<td>2.37459 x 10^9</td>
</tr>
<tr>
<td>Nearest Neighbors</td>
<td>447.505</td>
</tr>
<tr>
<td>Neural Networks</td>
<td>86.4512</td>
</tr>
<tr>
<td>Random Forest</td>
<td>288.083</td>
</tr>
<tr>
<td></td>
<td>272</td>
</tr>
<tr>
<td></td>
<td>713</td>
</tr>
<tr>
<td></td>
<td>181.5</td>
</tr>
<tr>
<td></td>
<td>262</td>
</tr>
<tr>
<td></td>
<td>171</td>
</tr>
</tbody>
</table>

The most reliable model turned out to be random forest [15], even if the numbers were still too far from the real ones. However, this discrepancy was due to the limited number of samples, according to the training parameters of Table 1.

Table 1. Parameters of the random forest training.

<table>
<thead>
<tr>
<th>Predictor Information</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input type</td>
<td>Image</td>
</tr>
<tr>
<td>Method</td>
<td>random forest</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>757.0 ± 2 x 10^2</td>
</tr>
<tr>
<td>Loss</td>
<td>8.39 ± 0.68</td>
</tr>
<tr>
<td>Evaluation time</td>
<td>61.8 µs/example</td>
</tr>
<tr>
<td>Predictor memory</td>
<td>116.0 kB</td>
</tr>
<tr>
<td>Training examples used</td>
<td>50 examples</td>
</tr>
<tr>
<td>Training time</td>
<td>26.8 s</td>
</tr>
</tbody>
</table>

5. Results and Discussion

The predictors were trained to “learn” images with information content linked to the spatial dispositions of purposes. By training the predictors on the vector representation of the configurations (e.g., by representing states as an array of numbers), we obtained poor results with almost all the predictors.

The failure of the linear regression model shows that the behavior of the automaton is profoundly non-linear: a very small change in state can cause a big change in the number of visitors, whereas the random forest model seemed to capture this non-linearity better than the other predictors. The dataset was in fact analysed as a set of weakly correlated subsets with different characteristics. The result was a sort of average on multiple decision trees, in the machine-learning sense.

The real states classified the spaces in terms of the number of visitors, as in the dendrogram of Figure 11.
The random forest partition shown in Figure 12 is apparently the most realistic.

The simulation and data processing identified the random forest as an effective method. It is precisely from this point of view that it is possible to apply statistical analysis at building scale. Not only this, but increasing application, at least in the types of building observed, is also foreseeable in the years to come.

The overall results indicate that the modeling of the Centquatre as a cellular automaton, together with finding the automaton’s rules by means of machine-learning tools, seems to be a promising way to build an instrument for economic planning. By monitoring the building continuously and collecting data on social and economic purpose and responses, the predictor was able to provide increasingly precise answers.

The predictions on the order of $10^9$ provided by the linear model show that the behavior of a complex building is not adequately represented by simple cause–effect proportionality. In addition, since some spatially distant activities may influence each other, the nearest neighbours model appears to be unsuitable. Finally, the unrealistic negative values returned by the neural networks model are probably a size effect which disappears with increasing input data. The random forest model seems the most reliable at efficiently mapping different structured choices.

The results achieved with a small volume of data were very encouraging, suggesting that it is worthwhile to replicate the approach with more extensive and robust data. Although deeper analysis
is required to explain the failure of some models, our procedure to select an optimal model suggests that the random forest model is an effective tool for evaluating the management of a complex building.

![Random Forest](image)

**Figure 12.** Random forest dendrogram.

It should be emphasised that a predictor is only a measurement instrument and does not replace space utilisation planning, which must be conducted on the basis of choices that take many variables into account. The 12 purposes we considered can be refined by considering the details of single activities. Within the limits of the available computational capabilities, this refinement should provide more precise and reliable predictions. Finally, as in the case of the Centquatre, behavior coding may also be required in other complex buildings, obviously in cases of multiple concurrent activities for which compatibility must be ensured.

To correlate the Centquatre configurations with the number of visitors, we represent its behavior as a multistate cellular automaton. Since each cell is related to the others in a complex way, the methods of linear regression and nearest neighbours gave incorrect outputs, while the long-range correlations were accounted by neural network [30] and random forest methods [15]. However the neural network method, which is quite efficient for image recognition, did not work in this case because the (open) problem of rule detection for a cellular automaton [31] is not an image recognition issue. On the other hand, the random forest, based on a decision trees method, was more effective. In fact, the resampling and averaging of the decision trees corresponds more closely to the “global” choices of the Centquatre administration.
6. Conclusion

Building on the results discussed in the previous sections, the complexity found at architecture scale in the Centquatre is typical of city complexity. In fact, we can take a step further and recognise our selected case study as a well-designed complex building, due to the marché de definition process and its architectural solutions. For this particular case, we discussed how the French approach has revolved around prefiguring a functional and spatial program for a single building (i.e., aiming to transform it into an incubator of uses). We therefore have a substantial degree of urban complexity in a single building. We also identified two dimensions that can help or enable this process: first, the application of big data-mining techniques to a building, and second, the essential value of indoor public space.

In fact, on one hand the behavior of a complex building can be described as a cellular automaton, and therefore analysed by machine-learning techniques. Indeed, the behavior of a complex building, as an incubator of uses like the Centquatre, can give rise to a variety of situations so numerous and varied that it is necessary to use big data-mining techniques—in particular those that exploit artificial intelligence. As the results indicate, this approach is viable with an appropriate classification model, such as random forest, and a sufficiently large and robust data set.

On the other hand, the present work shows that a complex building is the result of integrating public space in an architectural structure. As illustrated by the Centquatre, the large central area assigned as public space is the spatial component in which activities and uses are distributed in configurations (some examples are given in Figure 8). This is why the indoor public space is paramount, gathering different visitors in the Centquatre at the same time.

The relationship between the uses included in each configuration and the number of visitors exploiting the various uses is explored in the paper as a non-linear system, and is analysed using a random forest model. This relationship also shows that the presence of people in the indoor public space is a social and economic value. The presence of different people performing different activities in the same space is a form of aggregation proper to urban public space because it generates encounters and conviviality, even if it occurs in a building. Ordinary public life is incorporated in a building, which is an organised system with specific goals. Hence, this case study of the Centquatre shows that the presence of people can be considered an economic value in line with the goals of the system.

We have shown the possibility of the statistical-combinatorial approach to multifunctional building behavior, assimilating it to a multistate cellular automata. The identification of its updating rule can be done with the methods of machine learning related to random forest algorithm. However, since the random forest is more efficient when the number of trees is very large, it is necessary to apply this technique to a greater number of data belonging to different case studies.


Funding: This research received no external funding

Acknowledgments: The authors thank the subjects whose participation made this study possible. A. Mela, suggested to pursue this experimentation, J. Bourbon, provided the Centquatre-Paris data.

Conflicts of Interest: The authors declare no conflict of interest.

References

© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).