

<b>Project Number</b>	<b>ESPRIT / LTR /24 939</b>
<b>Project Title</b>	<b>AGENT</b>
<b>Deliverable Type</b>	<b>Report</b>

<b>Deliverable Number</b>	<b>D C4</b>
<b>Internal ID</b>	<i>AG-99-11</i>
<b>Contractual Date of Delivery</b>	<b>01/09/99 (technical Annex)</b>
<b>Actual Date of Delivery</b>	<b>22/09/99</b>
<b>Version</b>	<b>Final</b>
<b>Workpackage/Task contributing</b>	<b>C2</b>
<b>Authors</b>	<b>UNI-ED – IGN – UNI-ZH Principal author : N. Regnaud</b>
<b>Confidentiality</b>	<b>Public</b>

<b>Title of Deliverable</b>	Specifications for measures on MESO level & organisations
-----------------------------	---

<b>Abstract</b>	This report lists the contextual measures having already been developed in previous generalisation platforms and intended to be incorporated to the agent prototype to help in the contextual generalisation task. Measures are described through a common template which is explained in the introduction of the report.
<b>Keyword List</b>	Measures on organisation

## Table of content

<i>Table of content</i>	<u>2</u>
<i>1 Introduction</i>	<u>3</u>
<i>2 Characterisation of roads</i>	<u>4</u>
2.1 Characterisation of one road	<u>4</u>
2.2 Characterisation relations between roads	<u>8</u>
<i>3 Characterise buildings</i>	<u>12</u>
3.1 One building with regard to its neighbourhood	<u>12</u>
3.2 A group of buildings	<u>18</u>
<i>4 Conclusion</i>	<u>29</u>
<i>5 References</i>	<u>30</u>

# 1 Introduction

Contextual measures describe various properties that exist among a selection of objects. They can exist at a fine level of detail - such as the relative angle between the front of a building and the road, or these measures can describe the distribution pattern of a collection of objects - such as the variation of road network density from the city to the suburbs. The following pages list contextual measures deemed relevant to this phase of the AGENT project. It is by no means exhaustive. The list has evolved in response to information requirements of generalisation methods. For example a method that simplifies the network needs to know about the distribution of the network in order to preserve its homogeneity. Contextual measures that describe the topology, density, and connectivity of the network are therefore required.

The measures are described in a common template and are somewhat self descriptive. The use of a common template makes some of the fields redundant depending on the measure being described. By way of illustration, we describe the first measure 'relative orientation of a building to a road' in detail.

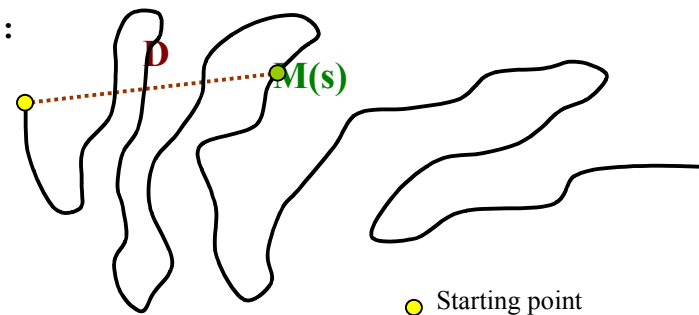
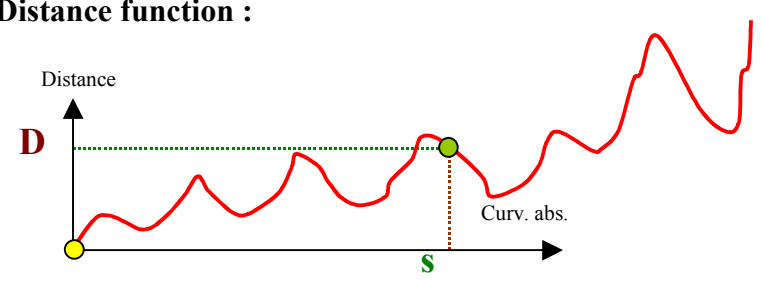
*Tool type* is 'measure'. The alternative tool type is 'ADS' when the tool computes or uses an Additional Data Structure. The *level* describes which agent the measure is associated with - micro, meso or macro level agents. *Location within the process* describes where in the generalisation process it is envisioned that the measure will be used. Almost all measures are used in the initial analysis phase (in order to characterise them) and as a post process (as part of the evaluation phase). By default *current use* tends to mirror the description given in *Location within the process* but can give more detail of methods that currently use this information. *Pre processing* describes other tools that are required to provide information prior to the operation of the method described. *Input data types* define the objects input to the method. *Concept* is a justification for the existence of the method. It describes what the tool gives as a result. *References* cite work in which the theory of the measure was explored. *Output data types* define the form of the output generated by the measure. *Parameter's significance* states whether there are known parameters such as tolerance values used by the measure, and explains the signification of their values. *Present state* describes whether the algorithm is at a theoretical stage of development, or whether it has been implemented within an experimental platform such as *Stratege* or *LaserScan* software. *Drawbacks, possible improvements, similar tools and remarks* are opportunities to clarify issues relating to the development of the algorithm such as similar algorithms, the degree to which they are generic in nature.

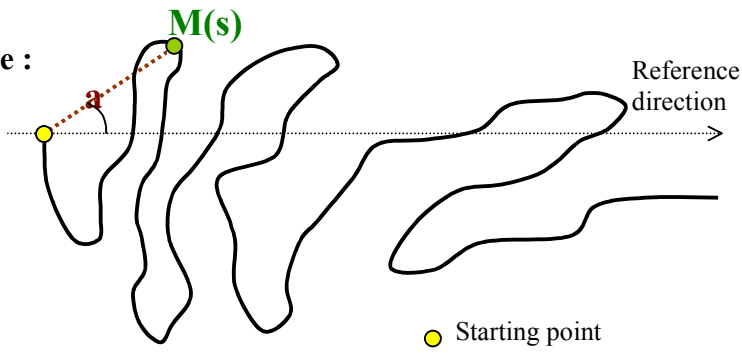
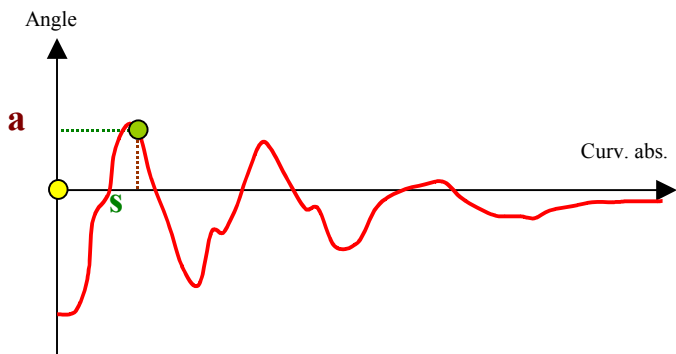
## 2 Characterisation of roads

Characterisation of roads have been largely described in C1 report. Here we report in a first section some ADS tools to represent lines in another way, to provide different information. The second section presents some contextual measures between linear features.

### 2.1 Characterisation of one road

Tool <i>type</i>	ADS
Tool <i>name</i>	<b>Distance function</b>
Level	MICRO/MESO
Location in the process	Intrinsic
	Before/after generalisation
	Characterisation/Segmentation
Pre-processing/tools required	-
Input data types	<b>Micro =&gt; linear</b>
	Road/River
Concept	It gives an idea of the spatial progression of the line (global shape, local and global regularity, ...)
Short description	We compute the euclidian distance between each point of the line and a starting point, usually an inflection point. So we obtain a function associating this distance to the curvilinear abscissa. We can represent it with a curve which is often globally increasing and possibly with maxima and minima.

	<p><b>1. Line :</b></p>  <p><b>2. Distance function :</b></p> 
References	(Barillot 99)
Output data types	Fonction
Parameters' significance	-
Present state	Implemented in ADA.
Drawbacks	-
Possible improvements	-
Similar tools	Angle function
Remarks	-

Tool type	ADS
Tool name	<b>Angle function</b>
Level	MICRO/MESO
Location in the process	Intrinsic
	Before/after generalisation
	Characterisation/Segmentation
Pre-processing/tools required	-
Input data types	<b>Micro =&gt; linear</b>
	Road/River
Concept	It gives an idea of the spatial progression of the line (global shape, local and global regularity, linearity, ...)
Short description	<p>We compute the angle between a reference direction and the segment joining a point of the line to a starting point, usually an inflection point. So we obtain a function associating this angle to the curvilinear abscissa. We can represent it with a schema very often with maxima and minima.</p> <p><b>1. Line :</b></p>  <p><b>2. Angle function :</b></p> 
References	Barillot 99

ESPRIT/LTR/24 939

Output data types	Fonction
Parameters' significance	-
Present state	Implemented in ADA.
Drawbacks	-
Possible improvements	-
Similar tools	Distance function
Remarks	-

## 2.2 Characterisation relations between roads

The three measures presented in this section address relative position of roads. The two first ones aim at characterising the proximity between roads, while the last one present a method for ordering proximity conflicts occurring in a set of roads

Tool type	Measure
Tool name	<b>Elongated proximity</b>
Level	MICRO-rel, MESO-rel
Location in the process	Intrinsic
	Before generalisation
	Characterisation
Current use	Before generalisation, to explicit the gravity of a conflict between two objects.
Pre-processing/tools required	ADS: Voronoï diagram between objects Algorithm : Conflict detection
Input data types	Voronoï edges and object boundaries
	separability threshold
Concept	The more continuous and narrow the free space strip between two objects, the more accute the conflict between them.
Short description	<p>The two objects are said to be in conflict when the minimal interdistance (read on the attributes of their common Voronoï interface) is shorter than <math>2.\sigma</math> (<math>\sigma</math> is a parameter and is equal to half the separability threshold).</p> <p>The portion of the Voronoï interface where the distance to the objects is smaller than <math>\sigma</math> is selected : the end-points are projected on the objects' contours, which delineates the conflicting area. <math>S(\sigma)</math> is this area's area, <math>L(\sigma)</math> is the length of the conflicting Voronoï interface.</p> <p>The conflict gravity indicator is : <math>S(\sigma) / L^2(\sigma)</math> (the smaller the value, the graver the conflict)</p>



References	(Hangouët 1998)
Output data types	Real
Parameters' significance	$\sigma$ : half the allowed separability threshold
Present state	described, programmable
Drawbacks	(none intrinsic)
Possible improvements	(none)
Similar tools	buffer overlap
Remarks	(none)

Tool <i>type</i>	Measure
Tool <i>name</i>	<b>Hausdorff distance</b>
Level	MICRO-rel, MACRO-rel
Location in the process	intrinsic, extrinsic
	Before or after generalisation
	Characterisation of the distance between two linear objects (or contours), or control of the distance between a generalised linear (or contour) object and its original version
Current use	After generalisation, to check whether the generalised version is not straying too far from the original.
Pre-processing/tools required	The Voronoï diagram of either object helps computing the Hausdorff distance
Input data types	Micros or mesos or mixed
	two objects
Concept	The Hausdorff distance between two objects expresses their positional remoteness
Short description	The Hausdorff distance (HDAB) between objects A and B involves two components: $D_{ab}$ and $D_{ba}$ , where $D_{xy}$ means the largest of the smallest distances from the points of X to Y. HDAD is the larger of the two components.
References	(Hangouët 1998)
Output data types	Real
Parameters' significance	no parameter
Present state	described, an approximation is now programmed in LAMPS2
Drawbacks	(none intrinsic)
Possible improvements	(none)
Similar tools	(none)
Remarks	(none)

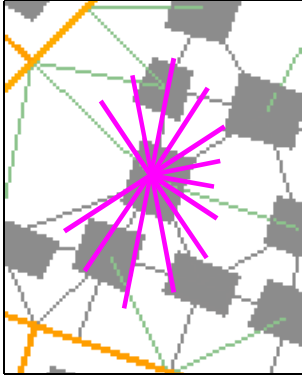
Tool type	Measure
Tool name	<b>Nickerson Ranking</b>
Level	MICRO-ext
Location in the process	Intrinsic
	Before generalisation
	Road symbol Conflict Ranking
Current use	Before generalisation, to select the worst overlap between a road and a neighbouring road.
Pre-processing/tools required	Algorithm : Access retrieval (i.e. neighbouring roads)
Input data types	The tested road
	Road network
Concept	Overlaps between road symbols are more or less complex to resolve. They then need to be ranked so as to solve first the most important.
Short description	Nickerson describes ten specific overlapping and rank them upon three parameters (common nodes, relative directions and length of the overlapping section). This ranking only handle small and unique intersections. So this list has been extended and modified so as to both integrate the configuration of an overlapping
References	Nickerson 1988
Output data types	Real value. The left part is the integer ranking. The right part helps in selecting the worst conflict with a same rank. It is based on the length of the overlapping.
Parameters' significance	<i>Epsilon</i> : Angle tolerance to define whether the two conflicting lines are perpendicular, parallel or not. <i>Extent</i> and <i>Tol</i> : Two parameters, used when an extreme node is involved, which define the threshold ( $Tol * Extent$ ) whether the length of the overlapping parts are small or not.
Present state	Implemented in Lamps2
Drawbacks	Nickerson's explanations remain obscure. Its ranking is useful for only a few cases among a generic road network. Nothing is mentioned about a polyline with multiple conflicts. Moreover, as soon as polyline reaches a significant extent, none of its ranking can still be applied.,
Possible improvements	Refine the definition of the relevant tangent.
Similar tools	An 'easy to implement' one should be, for long polylines, would a ranking based on : Number of conflicts along the lines, and size of them.
Remarks	Although this ranking has been used successfully by the COGIT staff, I still don't believe this <u>original</u> definition of the Nickerson ranking is really useful for a good generalisation purpose.

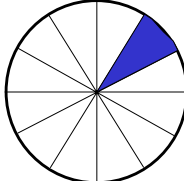
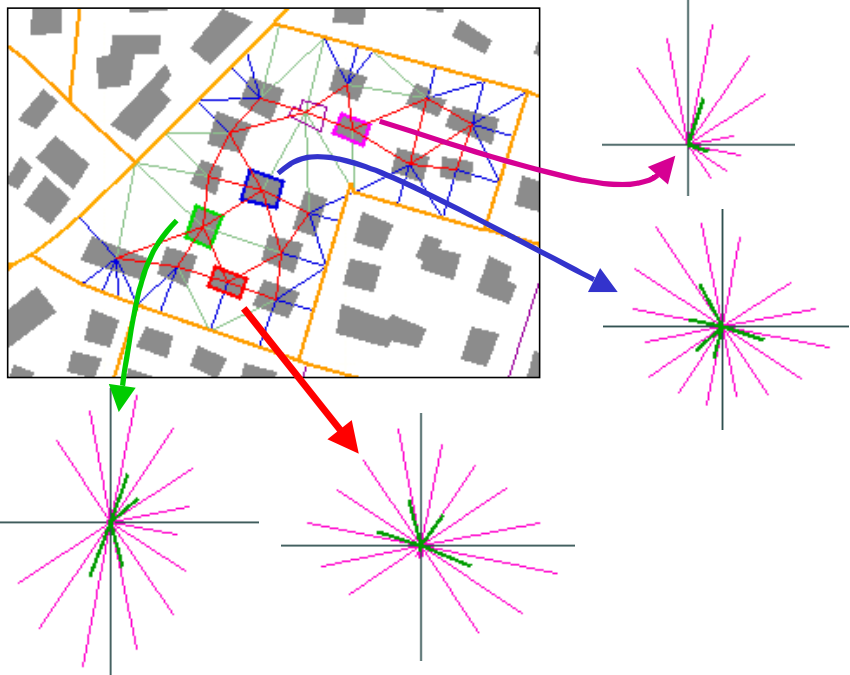
## 3 Characterise buildings

This chapter provides some tools and measures to describe buildings. We have split the chapter in two sections. First section describes tools for one building to evaluate its relationships with its neighbourhood. Second section present tools to characterise a set of buildings.

### 3.1 One building with regard to its neighbourhood

Tool <i>type</i>	Measure
Tool <i>name</i>	<b>relative_orientation</b> (of a building to a road)
Level	MICRO-ext
Location in the process	Intrinsic
	Before or after generalisation
	Characterisation
Current use	Before generalisation, to explicit the relations of a building to its access street.
Pre-processing/tools required	Algorithm : Access retrieval (i.e. association of building to road)
Input data types	<b>Meso =&gt;</b> Related micro agents
	building + street
Concept	The orientation of a building relative to the nearest street partakes to shaping urban or peri-urban landscapes.
Short description	The absolute orientation of the building is no longer expressed in the map coordinate system, but in a rotated version making the x-axis coincide with the tangent to the nearest street where the building's centroid projects.
References	(Hangouët 1998)
Output data types	Vector + quality indicator
Parameters' significance	no parameter
Present state	implemented in <i>Stratège</i>
Drawbacks	(none intrinsic)
Possible improvements	Refine the definition of the relevant tangent.
Similar tools	(none)
Remarks	(none)

Tool <i>type</i>	Measure/ Instantiated structure
Tool <i>name</i>	<b>Building-Congestion</b>
Level	MICRO
Location in the process	Intrinsic
	Before generalisation
	Characterisation
Pre-processing/tools required	ADS : PROXIMITY VALUED DELAUNAY TRIANGULATION
Input data types	<b>Micro within a Block...</b>
	Building
Concept	<p>The aim is to qualify the proximity of an object in different directions to evaluate its congestion.</p> <p>In such a way it is possible to compare object's congestion for different purposes such as autonomy order, building removal or displacement.</p> 
Short description	<ol style="list-style-type: none"> <li>1- The proximity between the building and its neighbours is computed from valued Delaunay triangulation.</li> <li>2- The smaller the proximity, the higher the congestion. Congestion vector is : <math>Cong = (\  dist-max - prox \  / dist-max ) Prox-direction</math> <ul style="list-style-type: none"> <li>▪ <i>Dist max</i> is the highest distance considered. It can come from the LDT or from distance threshold value.</li> <li>▪ <i>Prox</i> is the proximity quantity.</li> </ul> </li> <li>3- Congestion depends on direction <i>If an object has a close object in a direction, this direction and its neighbouring directions are congested</i></li> <li>4- To represent directional congestions a Rose Structure is used: A rose structure is composed of n parts:</li> </ol>

	<div style="text-align: center;">  </div> <p>5- Each part inherits congestion from the congestion vectors by propagation. Propagation is realised +/- 45 °from proximity direction.</p> <p>6- If a part inherit from different congestion vectors, the largest value is preserved. Thus, a congestion value within a part is between 0 and 1.</p> <div style="text-align: center;">  </div> <p>In Green: congestion vectors,                  In Purple: Congestion value within each part. <i>Values are emphasised for legibility purpose.</i> Some parts have no congestion vector.</p>
References	(Ruas 99) Page 172-175
Output data types	Two possible output: 1) <b>Instantiated structure</b> : each part of the rose has a congestion value 2) two global values: <ul style="list-style-type: none"> <li>• <b>average congestion</b> = <math>\Sigma</math> part-congestion / number of-parts</li> <li>• <b>Freedom ratio</b> = number of free part / number of part</li> </ul>
Parameters' significance	Three parameters are used: <ul style="list-style-type: none"> <li>- <i>dist-max</i>: to compare congestion either within a block or between blocks. Dist-max can be filter from the value Delaunay triangulation</li> <li>- <i>number of part of a rose</i>: default value = 16 parts</li> <li>- <i>angle of propagation</i> which must be coherent with the number of</li> </ul>

ESPRIT/LTR/24 939

	part. Default value is 45 °
Present state	Implemented onto Stratège (LISP + SMECI).
Current use?	Used for <i>Building removal</i> in case of over density, inside a cost function.
Drawbacks	<p>IF</p> <ul style="list-style-type: none"> <li>- <i>dist-max</i> is too large      OR</li> <li>- <i>number of part</i> too small    OR</li> <li>- <i>propagation</i> too large,</li> </ul> <p>THEN</p> <ul style="list-style-type: none"> <li>- all directions are congested</li> </ul>
Possible improvements	<p>1- The angle of propagation could depend on the distance (the greater the distance, the smaller the angle)</p> <p>2- Accurate congestion could be computed from solid angles between buildings, while integrating distance variation.</p>
Similar tools	Voronoi gardens (Hangouët 98) could provide close results.
Remarks	Could be used for displacement purposes.

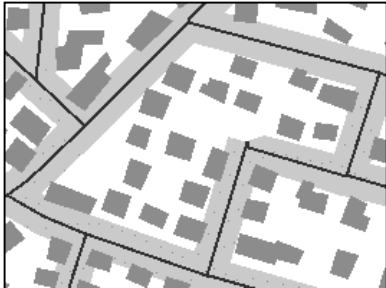
Tool <i>type</i>	Measure
Tool <i>name</i>	<b>Building median/deviation *- unusualness elicitation</b>
Level	MICRO-ext, MESO-int
Location in the process	intrinsic, extrinsic
	Before or after generalisation
	Decision on the unusualness of a building within a group of buildings
Current use	<p>Before generalisation, to detect buildings remarkable for the *-quality.</p> <p>The groups where it is applied are the Gestalt groups of buildings as described in (Regnauld 1998).</p> <p>Once a remarkable building is identified as remarkable, it is removed from the group, and further unusualness investigations are performed on the remaining group.</p>
Pre-processing/tools required	<p>(MST or any) grouping of buildings.</p> <p>Measurements of the *-quality for each building</p> <p>Measurements of the median and deviation values of the above measurements.</p>
Input data types	a group of buildings
Concept	<p>The *-quality might be of any type: for example, a building's own characteristic (nature, size, shape, height, orientation...), or one of its road-related characteristics (distance to the road, orientation to the road...), or one of its neighbouring buildings -related characteristics (interdistances between buildings...)</p> <p>A building with a *-quality extraordinary for the group it belongs to is remarkable, and may be important for the cartography desired. This is why it requires detection.</p>
Short description	The method checks whether the building's *-quality value lies within the (median- coef x deviation, median + coef x deviation) interval (coef is usually set to 2). If not, the building is marked as "extraordinary" for the *-quality.
References	Regnauld 98
Output data types	Semantic flag. – remarkable or not
Parameters' significance	no parameter
Present state	Programmed on <i>Stratège</i>
Drawbacks	(none intrinsic)
Possible improvements	(none)
Similar tools	<p>Building mean/deviation *- unusualness elicitation</p> <p>In this method, the median is preferred over the mean value because "an exceptional value in an homogeneous group is more distant from the median than from the mean value, and is thus more easily detected by this</p>



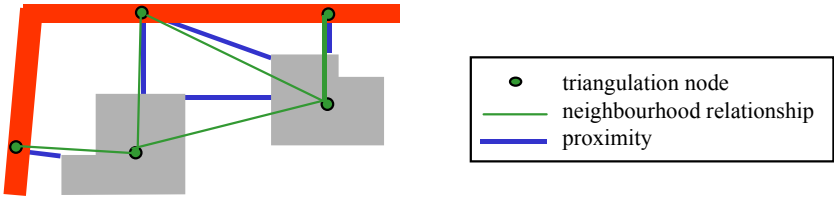
	method (Regnaud 1998 p.91)
Remarks	The remarkability measuring process is actually more elaborate. See (Regnaud 1998) pp.91 sqq., or in AGENT report D2 some of the commentaries to the Gestalt typification algorithm.

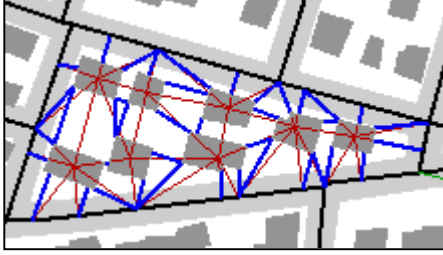
## 3.2 A group of buildings

This section regroups two types of tools. The two first measures describe an area containing some building. The description addresses the buildings inside and the limit of the area, while following measures address a group of buildings, with no references to the limits of the group. They are more generic.

Tool type	Measure
Tool name	<b>Block-density</b>
Level	MICRO-int
Location in the process	Intrinsic
	Before and after generalisation
	Characterisation and evaluation
Pre-processing/tools required	-
Input data types	City block
	Block component with their symbolisation width and their size goal
Concept	<p>The aim is to qualify the simulated density of a block while taking their default size goal into account</p>  <p><b>urban situation sit.33:</b>  <b>simulated density: 0,96</b>  elongation: 0,30  compactness: 0.86  average building size 148 m<sup>2</sup>  medium building size 152 m<sup>2</sup>  smallest building size 99 m<sup>2</sup></p>
Short description	<p>Given</p> <p>B a Block,</p> <p>S the situation</p> <p>B<sub>i</sub> a Building such that B<sub>i</sub> ⊂ S</p> <p>Area-m (B<sub>i</sub>) = max [Area (B<sub>i</sub>) ,minimum-building-size]</p> <p>R<sub>j</sub> a Road such that R<sub>j</sub> ⊂ Boundary (S)</p> <p>Area-s (R<sub>j</sub>) = road-symbol-width / 2</p> <p><b>Simulated_density (B) = Σ (Area-m (B<sub>i</sub>) + Area-s (R<sub>j</sub>)) / Area (B)</b></p>

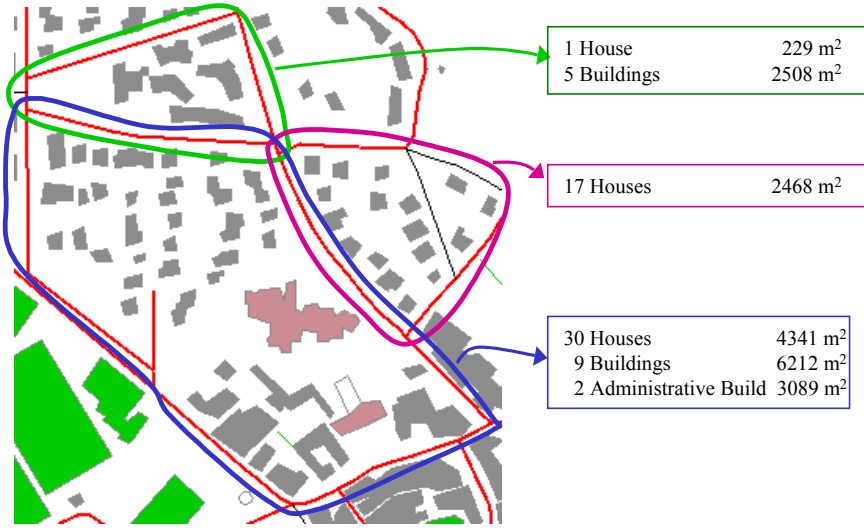
References	[Ruas 99] Page 119; [Ruas 98]
Output data types	Real value, should be between [0,1] But if buildings are very far from their goal size, the density is over 1.
Parameters' significance	Minimum-building-size: depends on the final scale Road-width depends on the symbology at the final scale.
Present state	Implemented onto Stratège (LISP + SMECI).
Current use?	Density is used to qualify a block constraint. It is mainly used to trigger object removal or typification
Drawbacks	-
Possible improvements	-
Similar tools	-
Remarks	-

Tool <i>type</i>	ADS: Instantiated structure
Tool <i>name</i>	<b>Block-Proximity</b>
Level	MESO-int
Location in the process	Intrinsic
	Before and after generalisation
	Characterisation and evaluation
Pre-processing/tools required	ADS : PROXIMITY VALUED DELAUNAY TRIANGULATION
Input data types	<b>Meso</b>
	Block
Concept	<p>The aim is to qualify the proximity between objects inside a block. According to the needs it can be buildings and streets or just building.</p> <p>Delaunay is used to define neighbourhood relationships:</p> <ul style="list-style-type: none"> <li>- An edge of the triangulation is a neighbourhood relationships</li> <li>- Proximity (distance) is an attribute of the edge.</li> </ul> 
Short description	<p>Delaunay triangulation connects a set of point with edges which never intersect. It creates a set of triangles whose angles are as close as possible.</p> <p>All the flexibility of Delaunay comes from the selection of points to connect, which depends on the needs. For our application the set of initial coordinates is not appropriate as two objects can be close without having close coordinates, and all coordinates can generate too many neighbourhood relationships.</p> <p>For our application we have chosen:</p> <ol style="list-style-type: none"> <li>1- The centre of building</li> <li>2- The projection of the centre on close lines</li> </ol>

	 <p>In case of complex line geometry some points on lines (such as nodes, convex hull or vertices) can be easily added</p> <p>Proximity (blue lines) can be computed with or without street symbolisation.</p>
References	[Ruas 99], Page 120-123
Output data types	<p>Two possible outputs:</p> <p>3) <b>Instantiated structure:</b> each edge of the Delaunay triangulation owns an attribut which describe the real distance</p> <p>4) <b>Distribution analysis of proximity:</b></p> <ul style="list-style-type: none"> <li>• Minimum, median, average, quartile 1 &amp; 3, maximum</li> </ul>
Parameters' significance	<p>Three parameters are used:</p> <ul style="list-style-type: none"> <li>- The function to select initial points <ul style="list-style-type: none"> <li>- Building centre</li> <li>- Projection on roads</li> <li>- Projection on lines</li> <li>- Inclusion of lines vertices, nodes and convex hull</li> <li>- Etc...</li> </ul> </li> <li>- The fact you introduce line's width or not</li> <li>- A distance can be used to reduce neighbourhood: <ul style="list-style-type: none"> <li>- IF real proximity is bigger than <i>dist-max</i></li> <li>- Then the neighbour relationships is inactive.</li> </ul> </li> </ul>
Present state	Implemented onto Stratège (LISP + SMECI).
Current use?	<p>Used for :</p> <ol style="list-style-type: none"> <li>1. Congestion computation</li> <li>2. Building removal</li> <li>3. Building displacement</li> <li>4. Block's proximity constraint</li> </ol>
Drawbacks	<ul style="list-style-type: none"> <li>- Computation time</li> <li>- Need a structure to be described (an edge is an object, a node also)</li> <li>- Possible redundancy in neighbourhood relationships between one building and the same street.</li> </ul>

ESPRIT/LTR/24 939

Possible improvements	Better coding ??
Similar tools	Voronoi [Hangouët 98] could provide better proximity description.
Remarks	Delaunay is not maintained during all the process.

Tool type	Measure
Tool name	<b>Semantic-Distribution-analysis</b>
Level	MESO-int
Location in the process	Intrinsic
	Before and after generalisation
	Characterisation and evaluation
Pre-processing/tools required	-
Input data types	<b>Meso</b>
	Block composed of buildings
Concept	The aim is to characterise the distribution of building type within a block
Short description	<p>The block classifies its buildings according to their type and compute their quantity in number and area.</p> <p>The classification can be done according to initial building type. To be more accurate, we differentiated houses and buildings according to a size threshold.</p> 
References	[Ruas 99] page 123
Output data types	List of (object-type, number, area) within a block
Parameters' significance	The threshold to distinguish Buildings and Houses
Present state	Implemented onto Stratège (LISP + SMECI).
Current use?	To compute the severity of semantic maintenance within a Block, specifically after buildings removal.
Drawbacks	! the houses which are dilated should not be classified as buildings !

ESPRIT/LTR/24 939

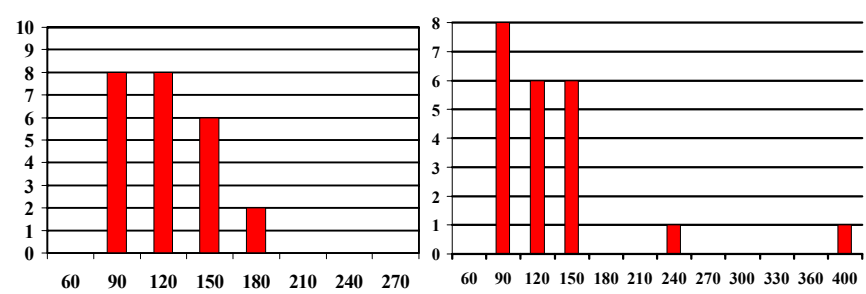
Possible improvements	Should be used to improve building selection
Similar tools	No
Remarks	-



Tool <i>type</i>	Characterisation
Tool <i>name</i>	<b>Homogeneity analysis</b> (of a set of values)
Level	MESO-int, + MICRO-ext
Location in the process	Intrinsic
	Before or after generalisation
	Decision on the homogeneity of the group, detection of possible exceptions
Current use	Usually applied on a group of buildings : either phenomenological groups of buildings accessible by a same street, on the same side of it, as described in (Hangouët 1998) for qualification or during the constitution of perceptual linear groups of buildings (Regnauld 98).
Pre-processing/tools required	(Phenomenological, or MST) grouping of buildings. Measurements of the quality for each building Measurements of the mean and deviation values of the above measurements.
Input data types	the <i>group of buildings</i> , with a measurement of the analysed quality for each of them
Concept	Buildings are qualified with regard to a criterion and the measures qualifies the group with regard to its homogeneity with regard to this criterion  The quality is of any type: it is for example, a building's own characteristic (nature, size, shape, height, orientation...), or one of its road-related characteristics (distance to the road, orientation to the road...), or one of its neighbouring buildings -related characteristics (interdistances between buildings...)  Assess the homogeneity of the quality over the group. If the group is homogeneous with some exceptional buildings, those ones are pointed out.
Short description	The method checks the “variation coefficient” of the values (standard deviation / median), and while less than a threshold (parameter = 0.9), removes the furthest value from the median, and starts again. At the end the remaining group is homogeneous, and if it represents more that a threshold percentage (parameter = 0.8) of the initial group, then the initial group is said homogeneous with the removed values as exceptions.
References	(Regnauld 1998)
Output data types	– Median – semantic flag (homogeneous or not), might be replaced by a real to give a more precise information – list of exceptional values
Parameters' significance	1 - tolerance for an homogeneous subgroup

ESPRIT/LTR/24 939

	2 - ratio of exceptions allowed in an homogeneous group
Present state	Implemented in <i>Stratège</i>
Drawbacks	In small groups, mean and deviation values are not statistically meaningful
Possible improvements	(none)
Similar tools	replace median by mean
Remarks	

Tool <i>type</i>	Measure
Tool <i>name</i>	<b>Distribution-analysis</b>
Level	MESO-int
Location in the process	Intrinsic
	Before and after generalisation
	Characterisation and evaluation
Pre-processing/tools required	Compute a set of value of the same character (e.g. proximity, size, density, etc.)
Input data types	<b>Meso</b>
	All
Concept	The aim is to characterise the distribution of a set of values e.g. building size, proximity; blocks density
Short description	<p>The characterisation of a distribution of values is shared into two parts:</p> <ol style="list-style-type: none"> <li>Statistical descriptors: <ul style="list-style-type: none"> <li>Minimum, average, quartiles (1, 2,3), maximum</li> <li>Standard deviation, variance, asymmetry, etc.</li> </ul> </li> <li>The detection of the distribution shape and particularities: <ul style="list-style-type: none"> <li>Shape of the distribution</li> <li>Clusters of values</li> <li>Exceptional values</li> </ul> </li> </ol> <p>Statistical parameters are very classical.</p> <p>Others, like the detection of clusters or exceptional value could be more complex and should be guided by constraints.</p> <p>For visual evaluation purposes a drawing of distribution by means of an histogram is useful:</p>  <p>The left histogram shows a distribution with peaks at 90 and 120, and a tail extending to 180. The right histogram shows a distribution with a peak at 90 and a tail extending to 400.</p>
References	[Ruas 99]

	<ul style="list-style-type: none"> <li>▪ Pages 124-131 for general description;</li> <li>▪ Pages 133-141 for the use of distribution for severity computation</li> </ul> <p>[Regnauld 98] Page 91: method for exceptional value detection</p>
Output data types	<ol style="list-style-type: none"> <li>1- Statistical descriptors</li> <li>2- Exceptional values</li> </ol>
Parameters' significance	<p>To compute an histogram, one can:</p> <ol style="list-style-type: none"> <li>1. Fix the number of classes : e.g. 20 classes</li> <li>2. Fix the quantity (the step) between two groups</li> <li>3. Filter the value (min, max) or (0, max) or (threshold 1, threshold2)</li> </ol> <p>This parameters allow to focus on some interesting values for the application. For example, if the need is to evaluate buildings size, big buildings are not interesting, too small buildings should not be in the same group then big enough ones.</p> <p>To detect exception, it can be necessary to give:</p> <ol style="list-style-type: none"> <li>1- What defines a group (from which number of entities)</li> <li>2- The required separation distance between a group and one isolated value. This value can depend on generalisation threshold.</li> </ol>
Present state	Implemented onto Stratège (LISP + SMECI).
Current use?	<ol style="list-style-type: none"> <li>1- To compute the severity of density, building size, proximity</li> <li>2- To change some component goals (e.g. building size)</li> </ol>
Drawbacks	Uses of statistical on small number of values => some statistical descriptors are not very significant.
Possible improvements	Need to develop specific functions for clusters and exception detection.
Similar tools	No
Remarks	To recognise some exceptional values the system needs the list of object identifier and value ((obj1,val1),...(objn,valn))

## 4 Conclusion

The template has attempted to formally describe the current list of proposed measures. It provides a basis for structured presentation of work undertaken as pre requisites to the generalisation process, and of importance in the evaluation phase. The creation of measures is driven by the input requirements of generalisation methods, which themselves arise from existing and on-going research. The development of such measures is deemed critical to the effective operation of agents in providing the framework and evaluation criteria for various decisions.

It is obvious that a lot more contextual measures will be needed for the agent prototype, but they have not been detected yet, as we need extensive tests on the first contextual prototype to detect where more control is needed, then which additional contextual measures are needed.

Some of these contextual measures depend on auxilliary data structures developed as part of the AGENT project. These are Voronoi, Delaunay, and Minimum Spanning Tree. These are described in this report under section 2.3.

## 5 References

(Barillot 99) Barillot, X. Généralisation automatique du linéaire : Etude de la courbure pour analyser la forme des routes. ACI'99 Ottawa, Canada, 1999.

(Hangouët 98) Hangouët, JF, Approche et méthodes pour l'automatisation de la généralisation cartographique ; application en bord de ville - PhD Thesis, Univ. De Marne La Vallée / IGN, 1998.

(Nickerson 88) Nickerson, B.G. Automated Cartographic Generalisation for Linear Features, Cartographica, vol. 25, No. 3, pp.15-66, 1988.

(Regnauld 98) Regnauld, N. : Generalisation du bâti: structure spatiale de type graph et représentation cartographique; PHd Thesis of Marne la Vallée University, 1998.

(Ruas 99) Ruas Modèle de généralisation de données géographiques à base de contraintes et d'autonomie. PhD Thesis, Univ. De Marne La Vallée / IGN, 1999.