ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

ΣΧΟΛΗ ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΓΕΩΓΡΑΦΙΑΣ

MICROSIMULATION MODELS IN GEOGRAPHY USING OBJECT-ORIENTED PROGRAMMING:

AN APPLICATION TO RESIDENTIAL MOBILITY

Υποδείγματα μικρο-προσομοίωσης στη Γεωγραφία με χρήση αντικειμενοστραφούς προγραμματισμού:

Εφαρμογή στη μετεγκατάσταση αστικών πληθυσμών

Marianthi Svinterikou

Advisory Committee:

Dr. P. Kanaroglou, Dr. N. Soulakellis, Dr. S. Koukoulas

2007

Acknowledgements

I would like to thank my supervisor Dr. Pavlos Kanaroglou for his guidance throughout this work. His vast experience on my research area was an important source of knowledge for me.

I would like to thank my colleague Justin Ryan at McMaster University in Hamilton, Canada, for his cooperation throughout this work. His useful comments improved the outcome of my research significantly.

I would like to thank Dr Mari Lavrendiadou who lectured at the Housing Geography course for her help in filling out my research questionnaires.

I am grateful to my beloved husband for his support the last years of this thesis. His patience and encouragement helped me overcome difficulties and make it all the way to the submission of this thesis.

I am grateful to my mother for her valuable help and support throughout this thesis. I would like to thank her for the countless hours of babysitting my new born daughter to help me make it all the way to the submission of this thesis.

I would like to thank my parents for their endless love and patience all these years and my beloved sister Sofi who stood by me during the years of my studies. Their support encouraged me to continue searching for knowledge.

I would like to thank my colleague and very good friend Makis Nikellis for helping me keep up a full time job with PhD research.

Στην κόρη μου Ιωάννα,
το σύζυγό μου Μιχάλη και
τη μητέρα μου Μαρία.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abstract

The need to mitigate the repercussions of urban sprawl over the last decades triggered the use of systematic urban plans that encompass land use, transportation and environmental dimensions. These plans aim at forecasting housing demand and analyzing the impacts governmental policy programs may have on urban development. While a considerable number of transportation models has been developed and used in the regional transportation planning process, review of the recent literature reveals that land use models have received less attention.

The lack of a fully documented and easily expandable land use model that incorporates formally the representation of spatial entities in the model triggered the need for the development of *URM-Microsim (Urban Residential Mobility Microsimulation)*; a *prototype decision-support system,* which captures the household development and the housing market changes of a city. Initially, the analysis part of *URM-Microsim* is defined which captures the requirements of the system to be developed. For the graphical representation of the analysis model, UML with spatio-temporal extensions is used in conjunction with the Microsoft Visio 2000 UML CASE tool. Next, the design model of *URM-Microsim* is presented taking into consideration the implementation platform to be used. Therefore, UML with ArcInfo extensions (ArcInfo Geodatabase Model) in conjunction with the Microsoft Visio 2000 UML CASE tool is used for the graphical representation of the design model. Then, *URM-Microsim* is implemented and applied in the real world. The Municipality of Mytilene is chosen as a testing ground. A population synthesis technique is used to produce the Mytilene micro-data at the individual and household level. Aggregate demographic and housing market data are gathered from various data sources (for example, the National Statistical Service of Greece, Eurostat,

etc). Furthermore, a survey questionnaire is defined and used to collect a small sample of household and dwelling micro-data for the city of Mytilene. In addition, the probability distributions required by *URM-Microsim* are calculated. Finally, *URM-Microsim* is loaded with the Mytilene data and the system's performance is examined; it is executed in order to simulate the population and housing market developments and the simulation results are discussed.

Περίληψη

Η άναρχη οικοδομική ανάπτυξη των πόλεων τις τελευταίες δεκαετίες προκάλεσε την ανάγκη ανάπτυξης μοντέλων που αφορούν τη χρήση γης, τις συγκοινωνίες και το περιβάλλον. Στόχος τους είναι να προβλέψουν την απαίτηση για κατοικία και να αναλύσουν τις επιδράσεις που τα κυβερνητικά πολιτικά προγράμματα μπορούν να έχουν στην αστική ανάπτυξη.

Η μελέτη της σχετικής βιβλιογραφίας έδειξε ότι ενώ ένας σημαντικός αριθμός μοντέλων έχει αναπτυχθεί και έχει χρησιμοποιηθεί κατά τη διαδικασία της περιφερειακής σχεδίασης των συγκοινωνιακών δομών διαφόρων πόλεων, λιγότερη πρόοδος έχει γίνει στον τομέα της χρήσης γης.

Ένας βασικός λόγος είναι το γεγονός ότι πρόκειται για συστήματα από τη φύση τους εξαιρετικά πολύπλοκα. Γι' αυτό το λόγο, κατά την ανάπτυξή τους είναι απαραίτητη η ομαδοποίηση των διεργασιών που περιλαμβάνουν, καθώς και η σταδιακή και αυξητική ανάπτυξή τους. Ακαδημαϊκοί και ερευνητές έχουν επικρίνει τα υπάρχοντα συστήματα για τα αμφίβολα αποτελέσματα τους και τα χαρακτηρίζουν ως ανεπαρκή για την πρόβλεψη και το σχεδιασμό της αστικής ανάπτυξης. Επιπλέον τα χαρακτηρίζουν ως «μαύρα κουτιά» θεωρώντας δύσκολο να κατανοήσουν τον τρόπο λειτουργία τους.

Όσο αφορά τα μοντέλα χρήσης γης, τα περισσότερα ασχολούνται με την χρήση γης για κατοικία, μία που αυτή παίζει τον καθοριστικό ρόλο στην οικιστική ανάπτυξη μιας πόλης. Για παράδειγμα στην περίπτωση του Δήμου Μυτιλήνης, το 81.43% των κτηρίων του διατίθενται για κατοικία, ενώ μόνο το 6% τους χρησιμοποιούνται για εμπορικούς σκοπούς. Επομένως, η οικιστική δομή μιας πόλης αλλάζει συνεχώς, κυρίως λόγω της προσφοράς

και της ζήτησης για κατοικία. Η οικιστική κινητικότητα σε μια πόλη εξαρτάται από μια σειρά αλληλένδετων αποφάσεων που λαμβάνονται σε επίπεδο ατόμων και νοικοκυριών, και εξαρτώνται από τις δημογραφικές, κοινωνικές και οικονομικές ιδιότητες των ατόμων και των νοικοκυριών (π.χ. οικογενειακή κατάσταση, εισόδημα και απασχόληση). Αυτές οι αποφάσεις προκαλούνται είτε από γεγονότα ανεξάρτητα των νοικοκυριών, όπως προγράμματα πολιτικής στέγασης και κυβερνητικές κοινωνικές πολιτικές, είτε από την αναμενόμενη δημογραφική ανάπτυξη των νοικοκυριών καθώς αυτά περνούν από τα διάφορα στάδια του κύκλου ζωής τους. Η απόφαση μπορεί επίσης να προκληθεί από φυσικές καταστροφές όπως ένας σεισμός.

Η γνώση των μελλοντικών αναγκών κατοικίας είναι μεγάλης σπουδαιότητας για όλους τους όσους συμμετέχουν στην αγορά κατοικίας όπως οι δημόσιοι και ιδιωτικοί επενδυτές, οι αρμόδιοι για το σχεδιασμό και οι κατασκευαστικές εταιρίες. Οι μελλοντικές ανάγκες κατοικίας και συνεπώς, η οικιστική κινητικότητα εξαρτώνται από τις αλλαγές στη δομή των νοικοκυριών και έχουν επιπτώσεις στη δυναμική της αγοράς κατοικίας. Επομένως, η ανάπτυξη των νοικοκυριών και οι αλλαγές στην αγοράς κατοικίας μπορούν να χρησιμοποιηθούν σε εφαρμογές όπως η πρόβλεψη της ζήτησης κατοικίας, καθώς επίσης και η ανάλυση των επιδράσεων που γεγονότα μπορεί να έχουν στους αστικούς πληθυσμούς και τη δομή των πόλεων.

Η μοντελοποίηση της οικιστικής κινητικότητας αποτελεί ένα ερευνητικό ζήτημα στις κοινωνικές επιστήμες εδώ και αρκετές δεκαετίες. Ο Rossi, ήταν ο πρώτος που το 1955 ανέλυσε τη σχέση μεταξύ των αλλαγών στον κύκλο ζωής ενός νοικοκυριού και της τάσης ενός νοικοκυριού να αλλάξει κατοικία.

Από τότε, ένας σημαντικός αριθμός μοντέλων έχει αναπτυχθεί και τα περισσότερα είναι μοντέλα συνάθροισης (η πληροφορία οργανώνεται και αναφέρεται σε σύνολα ατόμων της υπό μελέτη περιοχής). Ωστόσο, αυτά τα μοντέλα δεν είναι αποτελεσματικά λόγω της πολυπλοκότητας των δημογραφικών γεγονότων που βιώνουν τα άτομα και νοικοκυριά μιας πόλης. Επιπλέον, δεν μπορούν να μοντελοποιήσουν τη χωρική διάσταση της οικιστικής κινητικότητας του πληθυσμού.

Αυτά τα δύο μειονεκτήματα των μοντέλων συνάθροισης αντιμετωπίζει η μέθοδος της μικρο-προσομοίωσης (micro-simulation). Γι αυτό το λόγο ολοένα και περισσότερα μοντέλα βασίζονται στη μέθοδο της μικρο-προσομοίωσης (όπου η οργάνωση της πληροφορίας γίνεται σε επίπεδο κάθε ατόμου ξεχωριστά). Η έρευνα της σχετικής βιβλιογραφίας έδειξε ότι έχουν οριστεί οκτώ (8) τέτοια μοντέλα: HUDS, UPDATE, HOUSIM, LOCSIM, Fransson model, SVERIGE, ILUMASS και ILUTE. Από αυτά, τα έξι (6) πρώτα έχουν υλοποιηθεί. Η μελέτη και σύγκρισή τους έδειξε ότι αν και τα μοντέλα δεν λαμβάνουν υπόψη όλα τα στοιχεία και τις παραμέτρους της οικιστικής κινητικότητας, δύο (2) από αυτά, το LOCSIM και το μοντέλο του Fransson, έχουν την πιο ολοκληρωμένη ανάπτυξη και την πληρέστερη τεκμηρίωση. Και τα δύο έχουν αναπτυχθεί χρησιμοποιώντας εργαλεία πληροφορικής που βασίζονται σε παραδοσιακές τεχνικές ανάπτυξης συστημάτων. Επιπρόσθετα, δεν χρησιμοποιούν κάποια ολοκληρωμένη μέθοδο ανάπτυξης πληροφοριακών συστημάτων. Αυτό έχει δύο σημαντικά μειονεκτήματα:

- Δεν δίνουν τη δυνατότητα της χωρικής αναπαράστασης των αποτελεσμάτων του μοντέλου σε διάφορα επίπεδα συνάθροισης π.χ. σε επίπεδο γειτονιάς και οικοδομικού τετραγώνου.
- Είναι δύσκολη η βελτίωση και επέκτασή τους ιδιαίτερα όταν αυτό πρόκειται να γίνει από μία διαφορετική ομάδα ανάπτυξης από την

αρχική όπως συμβαίνει συνήθως σε μεγάλα και πολύπλοκα συστήματα όπως τα συστήματα οικιστικής κινητικότητας.

Αυτά τα δύο μειονεκτήματα αποσκοπεί να εξαλείψει αυτή η ερευνητική δουλειά. Πιο συγκεκριμένα, στόχος είναι ο καθορισμός και η ανάπτυξη του URM-Microsim (Urban Residential Mobility Microsimulation), ένα πρωτότυπο δυναμικό, χωρο-χρονικό σύστημα λήψης αποφάσεων, το οποίο προσομοιώνει την ανάπτυξη των νοικοκυριών και τις αλλαγές στην αγορά κατοικίας μιας πόλης χρησιμοποιώντας την τεχνική της μικρο-προσομοίωσης.

Για τον καθορισμό και την ανάπτυξη του URM-Microsim χρησιμοποιήθηκαν ισχυρά εργαλεία πληροφορικής τα οποία βασίζονται στην αντικειμενοστραφή τεχνολογία (την τελευταία τάση στην επιστήμη της πληροφορικής). Πιο συγκεκριμένα, έγινε επέκταση και συνδυασμός υπαρχόντων μεθόδων για τη δημιουργία μίας αντικειμενοστραφούς μεθοδολογίας η οποία διευκολύνει την ανάπτυξη χωρο-χρονικών συστημάτων λογισμικού. Έτσι λοιπόν:

- Για την ανάλυση και σχεδίαση χρησιμοποιήθηκε η ενοποιημένη γλώσσα μοντελοποίησης (UML) εμπλουτισμένη με χωρικές, χρονικές και χωροχρονικές έννοιες και μία διαδικασία Booch, προσαρμοσμένη στην ST-UML, η οποία περιλαμβάνει το σύνολο των βημάτων που καθοδηγούν την ανάπτυξη ενός συστήματος.
- Για την υλοποίηση του URM-Microsim χρησιμοποιήθηκε ένα γεωγραφικό σύστημα πληροφοριών, το ArcGIS, σε συνδυασμό με μια αντικειμενοστραφή γλώσσα προγραμματισμού, την Visual Basic.

Μετά τη μελέτη της υπάρχουσας βιβλιογραφίας η οποία πραγματοποιήθηκε στα πλαίσια του εννοιολογικού καθορισμού, πραγματοποιήθηκε η ανάλυση, ο σχεδιασμός και τελικά η υλοποίηση του URM-Microsim.

Κατά τη φάση της ανάλυσης καθορίστηκαν οι οντότητες που ενδιαφέρουν. Έτσι λοιπόν, ο πληθυσμός μιας πόλης αποτελείται από τα νοικοκυριά. Ένα νοικοκυριό έχει ένα ή περισσότερα μέλη. Κάθε μέλος είναι άτομο με μια συγκεκριμένη θέση στο νοικοκυριό (μέλος ζευγαριού, παιδί ή μόνος) και αυτό μπορεί να αλλάξει με την πάροδο του χρόνου. Ένα άτομο μπορεί να έχει ένα επάγγελμα δηλαδή μία θέση εργασίας (και επομένως, ένα μισθό) ή όχι και αυτό μπορεί να αλλάξει με την πάροδο του χρόνου. Η θέση εργασίας ανήκει σε μια περιοχή της πόλης (ζώνη κατοικιών). Ένα νοικοκυριό ζει σε μια κατοικία που νοικιάζει ή είναι ιδιόκτητη και που μπορεί να αλλάξει με την πάροδο του χρόνου. Μια κατοικία μπορεί να έχει έναν ή περισσότερους προγόνους (η κατοικία είναι η έκβαση μιας διάσπασης ή ενός συνδυασμού κατοικιών) ή έναν ή περισσότερους απογόνους (η κατοικία συμμετέχει σε συνδυασμό ή διάσπαση κατοικιών). Μια κατοικία ανήκει σε ένα κτήριο, το οποίο ανήκει σε ένα οικοδομικό τετράγωνο, το οποίο με τη σειρά του, ανήκει σε μια ζώνη κατοικιών (περιοχή). Η ζώνη κατοικιών έχει έναν ή περισσότερους γείτονες. Τα κτήρια, τα οικοδομικά τετράγωνα και οι ζώνες κατοικιών έχουν μια θέση στο χώρο, δηλαδή σαν αντικείμενα έχουν χωρική υπόσταση. Όλες αυτές οι οντότητες έχουν κάποια χαρακτηριστικά, οι τιμές των οποίων μπορούν να αλλάξουν με την πάροδο του χρόνου καθώς επίσης και οι σχέσεις μεταξύ των οντοτήτων. Για παράδειγμα, ένα άτομο έχει χαρακτηριστικά όπως το Φύλο, Ημερομηνία γέννησης και Οικογενειακή κατάσταση (ανύπαντρος, παντρεμένος, διαζευγμένος κ.λπ.). Η οικογενειακή κατάσταση ενός ατόμου μπορεί να αλλάξει με την πάροδο του χρόνου και αυτή η αλλαγή πρέπει να κρατηθεί στη βάση δεδομένων. Γι αυτό η οικογενειακή κατάσταση ορίζεται ως χρονικό χαρακτηριστικό.

Στη συνέχεια καθορίστηκε η συμπεριφορά των οντοτήτων του URM-Microsim. Πιο συγκεκριμένα, το σύστημα αποτελείται από τρία συστατικά μέρη (υπο-μοντέλα): τη *Ζήτηση κατοικίας*, την *Προσφορά κατοικίας*, και την *Αναζήτηση κατοικίας και μετακίνηση*. Η ζήτηση κατοικίας προσομοιώνει τα γεγονότα εκείνα που επηρεάζουν την απόφαση ενός νοικοκυριού να αναζητήσει μία νέα κατοικία (όπως δημογραφικά και οικονομικά γεγονότα). Η προσφορά κατοικίας προσομοιώνει τις αλλαγές στο απόθεμα κατοικιών και καθορίζει έτσι τον αριθμό κατοικιών που είναι διαθέσιμες στα νοικοκυριά. Τέλος, η αναζήτηση κατοικίας και μετακίνηση προσομοιώνει το ψάξιμο κατοικίας και την πραγματοποίηση των μετακινήσεων των νοικοκυριών που εξασφαλίζουν μία νέα κατοικία. Η έκβαση του επηρεάζει τόσο την προσφορά κατοικίας όσο και τη ζήτηση.

Η ανάγκη νέας κατοικίας εξαρτάται από διάφορους παράγοντες, όπως οι αλλαγές στη δομή του νοικοκυριού που προκαλείται από τα δημογραφικά γεγονότα (το θάνατο, τη γονιμότητα, το σχηματισμό ένωσης, τη διάλυση ένωσης, το σχηματισμό συγκατοίκων, τη διάλυση συγκατοίκησης, και τα παιδιά που φεύγουν από τους γονείς τους), τη μετανάστευση (έσω ή έξω μετανάστευση), τις αλλαγές στο εισόδημα και την απασχόληση του πληθυσμού, και την αλλαγή των απαιτήσεων για συγκεκριμένες υπηρεσίες κατοικίας (π.χ. δυνατότητα εύκολης πρόσβασης στη θέση εργασίας, ποιότητα γειτονιάς, κτλ). Καθώς άτομα και νοικοκυριά περνούν μέσα από τα διαφορετικά στάδια ζωής οι προτιμήσεις τους για κατοικία αλλάζουν, προκαλώντας την κινητικότητα. Όλα τα νοικοκυριά δεν έχουν την ίδια επιθυμία να βρουν μια νέα κατοικία. Η ένταση με την οποία ένα νοικοκυριό ψάχνει για μια νέα κατοικία εξαρτάται από το γεγονός που προκαλεί την επιθυμία για μετακίνηση. Για παράδειγμα, ένας διαζευγμένος επείγεται να βρει μια νέα κατοικία, ενώ μια οικογένεια με ένα νέο μέλος ψάχνει την αγορά κατοικίας με λιγότερη ένταση. Τρεις τύποι μετακινήσεων διακρίνονται με βάση τον προορισμό: νοικοκυριά που κινούνται μέσα στα όρια της αγοράς κατοικίας (οικιστική κινητικότητα), νοικοκυριά που κινούνται από την αγορά κατοικίας προς μία άλλη (έξω- μετανάστευση), και νοικοκυριά που κινούνται στην αγορά κατοικίας από τον υπόλοιπο κόσμο (έσω-μετανάστευση).

Κάθε άτομο υποβάλλεται στην πιθανότητα της βίωσης ενός από τα δημογραφικά γεγονότα καθώς και της εξω-μετανάστευσης, με βάση το τρέχον στάδιο του κύκλου ζωής του. Επιπλέον, καθορίζεται ο αριθμός των νοικοκυριών που εσω-μεταναστεύουν.

- *Θάνατος*. Κάθε άτομο υποβάλλεται στην πιθανότητα θανάτου με βάση την ηλικία και το φύλο του. Εάν το άτομο πεθάνει οι πληροφορίες σχετικά με το άτομο ενημερώνονται. Επιπλέον οι ιδιότητες του νοικοκυριού του ενημερώνονται καθώς επίσης και οι πληροφορίες σχετικά με τα υπόλοιπα μέλη του νοικοκυριού. Το νοικοκυριό εκφράζει την πρόθεση να μετακινηθεί σε μια άλλη κατοικία.
- *Γονιμότητα*. Κάθε γυναίκα του πληθυσμού υποβάλλεται στη πιθανότητα της γέννησης με βάση την ηλικία και την οικογενειακή κατάστασή της. Εάν η γυναίκα γεννήσει ένα παιδί, οι πληροφορίες σχετικά με το παιδί καθορίζονται και οι ιδιότητες του νοικοκυριού της μητέρας ενημερώνονται. Η οικογένεια εκφράζει την πρόθεση να κινηθεί προς μια άλλη κατοικία.
- *Σχηματισμός ένωσης*. Κάθε μη παντρεμένο άτομο ορισμένης ηλικίας αναζητά ένα μη παντρεμένο άτομο του αντίθετου φύλου και ορισμένης ηλικίας. Εάν το άτομο βρει ένα σύντροφο, το ζεύγος αποφασίζει να ζήσει μαζί (δημιουργία νοικοκυριού) και να ψάξει για μια άλλη κατοικία.
- *Διάλυση ένωσης*. Κάθε ζευγάρι που είναι παντρεμένο ή συμβιώνει υποβάλλεται στην πιθανότητα διάλυσης της ένωσης με βάση την ηλικία τους και τα έτη γάμου. Εάν το ζευγάρι χωρίσει, αποφασίζεται ποιος θα

εγκαταλείψει την κατοικία. Το άτομο που εγκαταλείπει την κατοικία δημιουργεί ένα νέο νοικοκυριό και αποφασίζει να ψάξει για μια άλλη κατοικία.

- *Παιδιά που φεύγουν από τους γονείς τους*. Κάθε άτομο της ηλικίας άνω των 18, που ζει ακόμα με τους γονείς του υποβάλλεται στην πιθανότητα να αφήσει τους γονείς του για να μείνει μόνο του. Εάν το άτομο αποφασίσει να αφήσει το πατρικό του σπίτι, δημιουργεί ένα νέο νοικοκυριό που πρέπει να ψάξει για μια άλλη κατοικία.
- *Σχηματισμός συγκατοίκων*. Ένα άτομο που ζει μόνο του ή συγκατοικεί με άλλους υποβάλλεται στην πιθανότητα εύρεσης ενός άλλου συγκατοίκου. Εάν το άτομο βρει έναν συγκάτοικο, αποφασίζουν να συγκατοικήσουν (δημιουργία νοικοκυριού) και να αναζητήσουν μια άλλη κατοικία.
- *Διάλυση συγκατοίκησης*. Ένα άτομο που συγκατοικεί με άλλα άτομα υποβάλλεται στην πιθανότητα της αναχώρησης από το νοικοκυριό για να ζήσει μόνο του. Εάν το άτομο αποφασίσει να αφήσει την αρχική κατοικία, δημιουργεί ένα νέο νοικοκυριό που πρέπει να ψάξει για μια άλλη κατοικία.
- *Έξω-μετανάστευση*. Ένα άτομο ή ένα νοικοκυριό μπορεί να αποφασίσει να έξω-μεταναστεύσει για διάφορους λόγους. Για παράδειγμα, ένα άτομο που είναι άνεργο για μεγάλο χρονικό διάστημα μπορεί να αποφασίσει να έξω-μεταναστεύσει. Ένα παιδί ηλικίας άνω των 18 μπορεί να αποφασίσει να σπουδάσει στο πανεπιστήμιο μιας άλλης πόλης. Εάν ολόκληρο το νοικοκυριό έξω-μεταναστεύει, η κατοικία του εγκαταλείπεται και προστίθεται στο σύνολο των διαθέσιμων κατοικιών.
- *Έσω-μετανάστευση*. Νοικοκυριά εκτός των ορίων της αγοράς κατοικίας αποφασίζουν να μετακινηθούν στην αγορά κατοικίας. Αυτά τα νοικοκυριά αποτελούνται από φοιτητές, ανώτερους υπαλλήλους του

στρατού, ξένους εργαζομένους κ.λπ. Αυτά τα νοικοκυριά πρέπει να βρουν μια κατοικία.

Κάθε άτομο ηλικίας 18 και άνω υποβάλλεται στην πιθανότητα να αλλάξει το εισόδημά του με βάση τις αλλαγές στην εκπαίδευσης του και τις επαγγελματικές τους αλλαγές.

- *Εκπαίδευση*. Το επίπεδο εκπαίδευσης του ατόμου καθορίζεται και το άτομο αποκτά μία σταδιοδρομία.
- *Είσοδος εργασίας*. Κάθε άτομο της ηλικίας 15 και άνω που δεν εργάζεται (οι στρατιώτες εξαιρούνται) υποβάλλεται στην πιθανότητα να βρει μια εργασία βασισμένη στο φύλο του ατόμου, την ηλικία καθώς και την τρέχουσα απασχόληση. Εάν το άτομο βρει μια εργασία, η οικογένειά του εκφράζει την πρόθεση της να μετακινηθεί σε μια άλλη κατοικία.
- *Αλλαγή εργασίας*. Κάθε άτομο, που εργάζεται ήδη, υποβάλλεται στην πιθανότητα να αλλάξει εργασία, να πάρει σύνταξη ή να γίνει άνεργος με βάση το φύλο του ατόμου, την ηλικία, καθώς και την τρέχουσα απασχόληση.

Ένα νοικοκυριό που δεν βιώνει κάποιο δημογραφικό ή οικονομικό γεγονός υποβάλλεται στην πιθανότητα της αλλαγής απαίτησης για συγκεκριμένες υπηρεσίες κατοικίας.

Η προσφορά κατοικίας τροποποιείται μακροπρόθεσμα προκειμένου να ικανοποιήσει τη ζήτηση κατοικίας. Οι αλλαγές στην προσφορά κατοικίας προκύπτουν από την κατασκευή των νέων κατοικιών, την κατεδάφιση και τη μετατροπή των υφιστάμενων κατοικιών (διαχωρισμός, συνδυασμός και ανακαίνιση κατοικιών) και τις αλλαγές στα έξοδα στέγασης.

- *Νέα κατασκευή*. Ο αριθμός των νέων κατοικιών που προστίθενται στο απόθεμα κατοικιών καθορίζεται με βάση την κινητικότητα της αγοράς και την έλλειψη κατοικίας. Η θέση κάθε νέας κατοικίας καθορίζεται με βάση την μη αναπτυγμένη γη που είναι κατάλληλη για την ανάπτυξη της περιοχής αγοράς και των τάσεων απαίτησης κατοικίας της αγοράς. Ο τύπος και οι υπηρεσίες της κατοικίας καθορίζονται με βάση τις τάσεις ζήτησης κατοικίας της αγοράς.
- *Μετατροπή κατοικίας*. Η δομή μιας κατοικίας μετατρέπεται σύμφωνα με κάποιο πιθανό τρόπο.
    - *Διάσπαση*. Κάθε κενή, μεγάλη κατοικία υποβάλλεται στην πιθανότητα διαχωρισμού σε δύο ή περισσότερες μικρότερες κατοικίες.
    - *Συνδυασμός*. Όλο το σύνολο δύο ή περισσότερων κενών, μικρών και γειτονικών κατοικιών υποβάλλεται στην πιθανότητα συνδυασμού σε μια μεγαλύτερη κατοικία.
    - *Ανακαίνιση*. Κάθε κατοικία σε μέτρια ή κακή κατάσταση υποβάλλεται στην πιθανότητα της ανακαίνισης. Εάν η ανακαίνιση πραγματοποιείται, ο τύπος της ανακαίνισης καθορίζεται (συντήρηση κατοικίας, επέκταση κατοικίας, αλλαγή αριθμού δωματίων κ.λπ.).
- *Κατεδάφιση*. Μια κατοικία σε κακή κατάσταση υποβάλλεται στην πιθανότητα της κατεδάφισης. Εάν η κατοικία είναι κατειλημμένη, το αντίστοιχο νοικοκυριό πρέπει να ψάξει για μια άλλη κατοικία. Η κατεδάφιση πραγματοποιείται μόνο μετά την εκκένωση της κατοικίας.
- *Αλλαγή στις δαπάνες στέγασης*. Οι δαπάνες στέγασης (κεφάλαιο συντήρησης, θέρμανσης, ηλεκτρική ενέργεια κ.λπ.) κάθε κατοικίας καθορίζονται με βάση τον τύπο της κατοικίας, το μέγεθος, την ηλικία και την ποιότητα κατασκευής.

Όσο αφορά την αναζήτηση κατοικίας και τη μετακίνηση. Όλα τα νοικοκυριά τα οποία επιθυμούν να μετακινηθούν ψάχνουν τη λίστα των διαθέσιμων κατοικιών για να βρουν μια κατοικία σύμφωνα με τους ακόλουθους κανόνες:

1. Αρχικά, το νοικοκυριό αποφασίζει εάν θα ψάξει για να αγοράσει ή να νοικιάσει μία κατοικία.
2. Το νοικοκυριό εξετάζει τις κατοικίες με βάση κάποια χαρακτηριστικά όπως η τιμή αγοράς/ ενοικίου, ο αριθμός δωματίων, ο τύπος κατοικίας, η ηλικία κτηρίου κτλ. Το νοικοκυριό καθορίζει τα χαρακτηριστικά που πρέπει να έχει η επιθυμητή κατοικία με βάση το γεγονός που βιώνει, το εισόδημα καθώς επίσης και αν πρόκειται για αγορά ή ενοικίαση κατοικίας. Ο αριθμός κατοικιών που εξετάζεται από το νοικοκυριό εξαρτάται από την ένταση αναζήτησης του νοικοκυριού. Με άλλα λόγια, τα νοικοκυριά με την υψηλή ένταση αναζήτησης εξετάζουν περισσότερες κατοικίες από αυτές με τη χαμηλή ένταση αναζήτησης. Εάν το νοικοκυριό βρει μία νέα κατοικία, τότε μετακομίζει εκεί και εκκενώνεται η τρέχουσα κατοικία του, διαφορετικά είτε παραμένει στην αρχική του κατοικία είτε εγκαταλείπει την πόλη (έξω-μετανάστευση). Μόλις πραγματοποιηθούν όλες οι μετακινήσεις, το ενοίκιο και η αγοραστική αξία των κατοικιών θα πρέπει να υπολογιστούν εκ νέου και να ενημερωθούν μια που οι τιμές τους είναι συνάρτηση της προσφοράς και της ζήτησης κατοικίας.

Στη συνέχεια πραγματοποιήθηκε η σχεδίαση τόσο της βάσης δεδομένων δηλαδή των οντοτήτων του URM-Microsim, όσο και της συμπεριφοράς τους. Διαγράμματα της UML χρησιμοποιήθηκαν για τον αναλυτικό σχεδιασμό της ροής των διαδικασιών του συστήματος και των γεγονότων που οι οντότητες βιώνουν.

Κατόπιν, πραγματοποιήθηκε η δημιουργία της βάσης δεδομένων χρησιμοποιώντας το Repository του Visio και στη συνέχεια τον Schema wizard του ArcCatalog. Κατόπιν προχωρήσαμε στην υλοποίηση των διαγραμμάτων της UML που αναπαριστούν τη συμπεριφορά των οντοτήτων της βάσης δεδομένων χρησιμοποιώντας τη Visual Basic for Applications του ArcGIS.

Προκειμένου να διαπιστωθεί η λειτουργικότητα του URM-Microsim εφαρμόστηκε στον Δήμο της Μυτιλήνης. Μια που δεδομένα σε επίπεδο ατόμου και νοικοκυριού δεν είναι διαθέσιμα από την ΕΣΥΕ, χρησιμοποιήθηκαν συνθετικά δεδομένα. Για την παραγωγή των δεδομένων αυτών χρησιμοποιήθηκε μία μέθοδο σύνθεσης πληθυσμού που ονομάζεται Συνδυαστική Μέθοδος Βελτιστοποίησης (Combinatorial Optimisation technique) ο αλγόριθμος της οποίας υλοποιήθηκε σε C++. Η μέθοδος δέχεται σαν είσοδο τα συνολικά στοιχεία του πληθυσμού, καθώς επίσης και δείγμα στοιχείων του πληθυσμού σε επίπεδο ατόμων και νοικοκυριών. Έτσι, για την περίπτωση του Δήμου Μυτιλήνης τα δεδομένα εισόδου που χρησιμοποιήθηκαν είναι:

- Γεωγραφικά δεδομένα και συγκεκριμένα τα οικοδομικά τετράγωνα με κωδικούς ΕΣΥΕ καθώς και τα κτίρια ανά οικοδομικό τετράγωνο από το Εθνικό Κτηματολόγιο.
- Συνολικά περιγραφικά δεδομένα από την Απογραφή 2001 της ΕΣΥΕ και ιδιαίτερα σημαντικά ήταν ο αριθμός νοικοκυριών ανά οικοδομικό τετράγωνο .
- Μίκρο περιγραφικά δεδομένα, δηλαδή δεδομένα σε επίπεδο ατόμου και νοικοκυριού τα οποία συλλέχθηκαν με τη χρήση ερωτηματολογίου.

Το αποτέλεσμα του προγράμματος σύνθεσης είναι τέσσερις πίνακες με τα επιθυμητά δεδομένα δηλαδή τον συνθετικό πληθυσμό του Δήμου

Μυτιλήνης, τα οποία προκειμένου να εισαχθούν στη βάση δεδομένων υλοποιήθηκαν οι κατάλληλες ρουτίνες σε Visual Basic.

Επιπλέον, για τον υπολογισμό των παραμέτρων του συστήματος (όπως η πιθανότητα θανάτου, γέννησης κτλ) χρησιμοποιήθηκαν αθροιστικά δεδομένα της EUROSTAT.

Στη συνέχεια, πραγματοποιήθηκαν κάποιες προσομοιώσεις για το έτος 2002 προκειμένου να διαπιστωθεί η λειτουργικότητα του συστήματος. Αρχικά, μία βασική προσομοίωση πραγματοποιήθηκε χρησιμοποιώντας τον συνθετικό πληθυσμό και τις τιμές των παραμέτρων του συστήματος για το Δήμο Μυτιλήνης. Στις ακόλουθες προσομοιώσεις, προκειμένου να διαπιστωθεί η επίδραση των δημογραφικών γεγονότων στην μετακίνηση του πληθυσμού, διατηρήθηκαν σταθερές όλες οι παράμετροι εκτός από μία κάθε φορά. Πιο συγκεκριμένα, σε κάθε προσομοίωση διπλασιάστηκε η πιθανότητα του παιδιού που φεύγει από τους γονείς του, της δημιουργίας ένωσης και της διάλυσης μιας ένωσης, αντίστοιχα.

Ο μεγάλος ανασταλτικός παράγοντας στην ανάπτυξη του URM-Microsim ήταν η έλλειψη δεδομένων για την πόλη της Μυτιλήνης. Για πολλά γεγονότα τα δεδομένα είτε ήταν ελλιπή, είτε δεν ήταν διαθέσιμα.

Καταρχήν, δεν υλοποιήθηκε η προσφοράς κατοικίας. Στο URM-Microsim όλες οι νέες κατοικίες και οι μετατροπές κατοικιών δίνονται εξωγενώς. Το ίδιο γίνεται και με τις αλλαγές στις τιμές των ενοικίων και των αγοραστικών αξιών των κατοικιών. Ακόμη το γεγονός αλλαγής εισοδήματος το οποίο συμπεριλαμβάνεται στην ανάλυση και το σχεδιασμό του URM-Microsim δεν υλοποιήθηκε. Έτσι, τα νοικοκυριά δεν ψάχνουν για νέα κατοικία λόγω αλλαγής στο εισόδημά τους. Επιπλέον, κατά την εύρεση εργασίας δεν λαμβάνεται υπόψη η χωρική διάσταση της νέας εργασίας (περιοχή στην οποία βρίσκεται η εργασία). Τέλος όσο αφορά την εσω-μετανάστευση, λαμβάνονται υπόψη μόνο οι φοιτητές που έρχονται κάθε χρόνο στη Μυτιλήνη για σπουδές, ενώ υπάρχουν και άλλες κατηγορίες νοικοκυριών που μετακομίζουν στην πόλη της Μυτιλήνης. Τέλος, το σύστημα εξετάστηκε και αξιολογήθηκε με την πραγματοποίηση ενός συνόλου προσομοιώσεων για το έτος 2002. Ωστόσο, μία πλήρης αξιολόγηση και αξιοποίηση του URM-Microsim δεν ήταν εφικτή μια που δεν ήταν διαθέσιμα εμπειρικά δεδομένα για την πόλη της Μυτιλήνης

URM-Microsim είναι ένα πρωτότυπο σύστημα και σαν τέτοιο, επιδέχεται πλήθος βελτιώσεων και επεκτάσεων.

Καταρχήν, κάποια από τα γεγονότα που υλοποιούνται στο URM-Microsim μπορούν να βελτιωθούν περαιτέρω όπως η αλλαγή στο εισόδημα ενός νοικοκυριού, η εύρεση εργασίας λαμβάνοντας υπόψη τις κενές θέσεις εργασίας και τους χώρους εργασίας, και η εσω-μετανάστευση. Επιπρόσθετα, ο αλγόριθμος αναζήτησης κατοικιών πρέπει να διαμορφωθεί για να ενσωματώσει όλες τις διαστάσεις αναζήτησης. Αυτή τη στιγμή οι σημαντικότερες διαστάσεις αναζήτησης εξετάζονται, δηλαδή ο αριθμός δωματίων, η τιμή, ο τύπος και κατάσταση ενοικίασης.

Η υλοποίηση της προσφοράς κατοικίας. Ιδιαίτερα, στατιστικά μοντέλα θα πρέπει να χρησιμοποιηθούν και να ενσωματωθούν στο σύστημα για τον υπολογισμό των τιμών των ενοικίων και της αγοραστικής αξίας των κατοικιών σαν συνάρτηση της προσφοράς και της ζήτησης κατοικίας.

Η εφαρμογή του URM-Microsim σε μία άλλη πόλη με περισσότερα δεδομένα θα δώσει τη δυνατότητα εξαγωγής περισσότερων συμπερασμάτων για την μετεγκατάσταση των νοικοκυριών της πόλης και καλύτερης αξιοποίησης των δυνατοτήτων του συστήματος.

Τέλος ο σχεδιασμός και η ανάπτυξη ενός μοντέλου μεταφορών ως επέκταση του URM-Microsim, θα δώσει τη δυνατότητα ανάπτυξης ενός ολοκληρωμένου συστήματος μεταφορών και χρήσης γης.

Chapter 1
Introduction

1.1 Introduction

The continuous waves of urbanization in most countries over the last few decades, have led to urban sprawl and car dependency with unfavourable repercussions for the quality of life in urban areas. Population swelling in cities and sprawl have been linked to excessive congestion, environmental problems with adverse effects on human health, and economic and fiscal problems in maintaining the city infrastructure. The need to mitigate these repercussions triggered the need of developing systematic urban plans that encompass land use, transportation, environmental and economic dimensions. A basic need within such plans is the ability to forecast housing demand and supply and to analyze the impacts governmental policy programs may have on urban housing markets. Unfortunately, the data analysis tools available to planners and policy makers have lagged far behind the need for developing coherent policy programs, leaving decision makers inadequately prepared to anticipate the potential consequences of alternative policy initiatives or infrastructure choices.

Until recently, another inhibiting factor in developing such tools has been the hardware limitations in computer technology. In many cases the use of data analysis and modelling techniques lead to rather contradictory or at least poorly integrated policy objectives (Waddell, 1998). Furthermore, the market effects and constraints of policy initiatives are given insufficient attention or are misunderstood. Therefore, the use of effective data analysis techniques as an objective for urban development planning is not in question.

Over the last three decades, while a considerable number of transportation models has been developed and used in the regional transportation planning process of several cities, especially in North America, land use models received substantial criticism for their poor performance (Waddell, 1998). It is striking that all integrated land-use and transportation models such as ILUTE (Miller *et al.*, 2004), ILUMASS (Moeckel *et al.*, 2002) and MOSES[1] (Birkin *et al.*, 2005; Birkin *et al.*, 2006) pay considerable attention to transportation modeling marginitizing the treatment of land-use modeling. Moreover, the field of modeling land use and transportation systems has been consistently criticized for its complexity and black box character (Timmermans, 2003). It is within this context that this work focuses on, modeling land use.

Among competing land uses within a city, residential land use is perhaps the most significant in determining the urban structure. For example, in the case of the Municipality of Mytilene, 81.43% of its buildings are allocated to residential use, while only 6% of them can be characterized as commercial. Therefore, the structure of a city constantly changes, mainly due to *housing demand* and *housing supply*. Housing demand depends on several factors, such as changes in household structure caused by *demographic events* (for example mortality, fertility, marriage/divorce etc), *migration* (in- or out-migration), *changes in income and employment* of the urban population, and *demand for specific housing qualities* (for example accessibility to employment location, structure type, neighborhood quality). Housing

[1] The MOSES system has not been implemented yet.

supply, on the other hand, aims to satisfy housing demand. Urban housing markets are, thus, constantly adjusting to new demand/supply equilibria.

Residential mobility in a city is highly dependent on a series of *interrelated decisions* that are taken both at *individual and household levels*, and depend on the demographic, social and economic attributes of individuals and households (for example marital status, income and employment). These decisions are triggered either by *shock events* external to the households such as housing policy programs and government social policies or by the expected household and family *demographic development* as they progress in their life cycle. Decision may also be triggered by natural disasters such as an earthquake.

Knowledge of future housing needs is of profound importance to all actors involved in the market such as public and private investors, planners and housing corporations. Future housing needs and consequently, residential mobility depends on changes in household structures and affects housing market dynamics. Therefore, modeling household development and housing market changes may be used in applications such as forecasting housing demand, as well as analyzing the impacts that shock events may have on urban populations and city structure. These models result in complex behavioral systems and their development, management and exploitation require the use of *effective development tools*.

Conceptualizing urban residential mobility has been a long-standing research issue in the social sciences (Rossi, 1955; Brown and Moore, 1970). Urban academics and policy makers have used over the years various *data analysis and modeling techniques* for the development of residential mobility models. While some of them do not produce reliable representations of these complex systems, others deal with this complexity more effectively (Bartlema and Vossen, 1988). The field of *computer science* has defined *several methodologies* for the development of software systems. However, not all of them handle the requirements of complex systems and thus, are not suited for residential mobility systems (Booch, 1994). The choice of effective *data analysis techniques* and *powerful computer science tools* is an important issue in developing residential mobility systems.

1.2 Objectives of the investigation

The aim of this work is the design and development of a *prototype decision-support system,* called *URM-Microsim (Urban Residential Mobility Microsimulation)*, which captures the household development and the housing market changes in a city. Its development will be shaped by several constraints that emanate from the context described in the previous section. First, *URM-Microsim* should be based on sound theory, offering the ability to analyze the impact of the events experienced by individuals and households on population residential mobility and consequently in the change of a city structure. Second, it should be defined in a way that it would be comprehensible to all actors involved in the development of such a system and particularly to the system's end users i.e. the policy makers. Third, it should be defined and then implemented in a way that it would be easily extended with additional functionality even by a team different than the original development team. Fourth, it should be defined and implemented in a way that it would make it adaptable to cities of varying size and complexity. Finally, it should be developed and tested within one city, using existing data.

In order to accomplish the above objectives:

1. An effective data analysis technique called *microsimulation* will be used to define the *URM-Microsim* model.
2. A computer science modeling language, called the *Unified Modeling Language (UML)* enriched with *spatial, temporal* and *spatio – temporal concepts* will be used for the graphical representation of the *URM-Microsim* model, during the analysis and design phases of the system development life cycle.
3. A powerful computer science tool i.e. the *ArcGIS* platform, which is a Geographic Information System (GIS) tool, will be used in conjunction with Visual Basic for Applications (VBA), an object-oriented programming language, as the system's implementation platform.
4. The use of the developed system will be demonstrated by applying it to the city of Mytilene. For this purpose *demographic and housing market data* for the city of Mytilene will be gathered from various data sources (for example, National Statistical Service of Greece, Eurostat, Surveys). Using these data a synthetic population of individuals and households will be created. The synthetic population will be inserted in the developed system, in order to simulate the residential mobility in the town of Mytilene, on a yearly basis.

1.3 The structure of the thesis

The development of a software system such as *URM-Microsim* consists of a set of phases that follow each other seamlessly and constitute the *system development life cycle*. The thesis is structured according to these phases, which are (Booch, 1994):

Conceptualization: the scope of the system is defined and the core requirements for the system to be developed are identified. Earlier in this chapter (**Chapter 1)** we addressed the issues that need to beincluded in the phase of conceptualization.

Analysis: the function of the system is defined, that is, what the system does, through domain analysis and use case planning. In *Domain analysis* existing related systems are examined in order to identify the classes and objects that are common to all applications within a given domain, such as housing market systems. The issues included in the domain analysis are addressed in Chapter 2 and Chapter 3. In particular, **Chapter 2** discusses the research conducted on the residential mobility process and examines existing residential mobility microsimulation models. Initially, the factors thought to influence residential mobility are discussed, as well as the most significant results that have been accumulated over the last forty years. Then, the different strategies employed in the modeling of the residential mobility process are presented. Next, the microsimulation technique, as well as its strengths is presented and the key concepts of a residential mobility microsimulation model are defined. Following that, we describe LOCSIM at length. This seems to be the most well developed and documented microsimulation model of residential mobility. Next, the HUDS, Fransson, SVERIGE, ILUMASS, UPDATE, HOUSIM and ILUTE residential mobility models are examined, and their main strengths and weaknesses are highlighted. Finally, an overview of the methods presented is given and the contribution of this work in the field of residential mobility systems is outlined.

Chapter 3 discusses powerful computer science tools i.e. models, methods and processes which allow the effective conceptual and graphical representation of a system under development such as a residential mobility system. Initially, the kinds of system development methods are presented where the predominance of object-oriented technology is pointed out. Next, the elements of object-oriented methods are presented. Then, the Unified Modeling Language (UML) is presented, which is defined by the Object Management Group (OMG), as *the standard object-oriented* modeling language and the use of UML in microsimulation model development is illustrated, through an example: applying UML to the existing model LOCSIM. In relation to **Chapter 3, Appendix A** provides the detailed graphical notation of the elements used in UML diagrams. The remaining sections of chapter 3 discuss the spatial, temporal and spatiotemporal concepts of systems and examine two existing Spatio-Temporal modeling languages based on UML called Extended SpatioTemporal UML and SpatioTemporal PVL for UML as well as the ArcInfo Geodatabase Model. The main strengths and weaknesses of these models are pointed out. Moreover, an overview of the modeling languages presented is given and the choices made within the context of this work are justified. Finally, the ST-UML process is defined, that is, a set of activities that guide the development of a system, by applying the UML modeling techniques extended with spatiotemporal constructs. In **Appendix B** the SpatioTemporal UML (STUML) Process is described in detail.

Use case planning: once domain analysis is completed, the analysis phase moves on to use case planning in order to develop a model of the system's desired behavior. Use case planning is addressed in **Chapter 4** where the analysis model of *URM-Microsim* is presented. The analysis model captures the requirements of the system to be developed and does not contain any implementation details in order to be usable by all members of a development team (users, analysts, programmers etc). For the graphical representation of the analysis model, UML with spatio-temporal extensions is used in conjunction with the Microsoft Visio 2000 UML CASE tool. Initially, the behavior of the system to be developed (what the system does) is presented that is, the *URM-Microsim Model* as well as its components, namely, *the Housing Demand, Housing Supply* and *Residential Search and Migration* sub-models. The *use-cases scenarios* that define in detail the events modeled by *URM-Microsim* are presented in **Appendix C**. Then, the *URM-Microsim* Database is presented, namely, the *entities* that are of interest as well as their *characteristics* and the *relationships* with each other.

Design: the architecture of the system under development is defined. Design is addressed in **Chapter 5** where the design model of *URM-Microsim* is presented taking into consideration the implementation platform to be used. Therefore, UML with ArcInfo extensions (ArcInfo Geodatabase Model) in conjunction with the Microsoft Visio 2000 UML CASE tool is used for the graphical representation of the design model. Initially, the behavior of the *URM-Microsim Model* and its sub-models *(Housing Demand, Housing Supply* and *Residential Search and Migration)* is presented, that is, the steps the model goes through, in order to execute the model processes. Next, the events that occur in the Housing Demand and the Residential Search and Migration sub-models are presented. In addition, the *URM-Microsim* Database is presented, namely, the *entities* that are of interest as well as *their characteristics* and the *relationships* with each other. Finally, implementation issues concerning the implementation of the *URM-Microsim* design model are discussed.

Evolution: the implementation is evolved though successive refinement in a series of releases. *URM-Microsim* is implemented and **Chapter 6** refers to the last system implementation releases. Specifically, the chapter is devoted to applying *URM-Microsim* to the real world. The Municipality of Mytilene was chosen as a testing ground for *URM-Microsim*. Mytilene was chosen for two practival reasons. First, it is a small town (about 30.000 population) with a single economy and quite self-contained. Secondly, since it is the seat of the Department of Geography of the University of the Aegean, it was reasoned that it would be easier to collect the required data and in cases where no data are available, more realistic assumption could be made. Initially, a short introduction on the history and background of Mytilene, its population and its housing market are presented. Next, the population synthesis technique used to produce the Mytilene data is presented, as well as its input data and an excerpt of the synthetic data is given. The questionnaire used to collect household and dwelling information for the city of Mytilene is presented in **Appendix D**. Finally, the probability distributions required by *URM-Microsim* are calculated.

Once *URM-Microsim* is loaded with the Mytilene data, the system's performance is examined in **Chapter 7**. The system is executed in order to simulate the population and housing market developments. Then, the issue of model efficiency arises: how well does the model simulate reality? Initially, the notion of system calibration is discussed, then the series of simulations carried out for *URM-Microsim* are presented and the simulation results are discussed. Next, the notion and methods of system validation are discussed.

Finally, in **Chapter 8** an overview of the work presented in this thesis, a summary of the main contributions of the research, and its conclusions are presented. In addition, further work in this area is discussed.

Chapter 2
Background Literature
in Modeling Residential Mobility

2.1 Introduction

In this chapter, we discuss the residential mobility research, including theories and empirical findings, as well as the relevant microsimulation models. Section 2.4 presents the factors thought to influence residential mobility and the most significant research findings over the last forty years. Section 2.5 presents the different strategies employed in the modeling of the residential mobility process. Section 2.6 presents the microsimulation technique along with its strengths. Section 2.7 defines the key concepts of a residential mobility microsimulation model. Section 2.6 gives an extended description of the LOCSIM model, which seems to be the most well developed and documented microsimulation model with respect to modeling demographic processes. Sections 2.7 to 2.13 examine the HUDS, Fransson, SVERIGE, UPDATE, HOUSIM, ILUTE and ILUMASS residential mobility models, and their main strengths and weaknesses are highlighted. Finally, section 2.14 gives an overview of the models presented and the contribution of this work in the field of residential mobility systems is outlined.

2.2 Residential mobility

Residential mobility can be defined as the movement of individual households from one place of residence to another. Regarding intra-urban mobility, these movements take place within the confines of a particular urban area. This form of population movement is quite substantial in any given year. On average, approximately 20% of all North American households move each year and about two-thirds of these residential shifts are intra-urban in nature (Short, 1978; Knox, 1987). This high level of movement indicates that the intra-urban migration process has an enormous potential to alter spatial relationships within urban areas.

Given this ability for population redistribution, it is not surprising that residential mobility has become a major area of interest in geographical research. This interest was sparked by a major paradigm shift within the field of geography. From approximately 1930 to 1960, areal differentiation was the major method of geographic inquiry. During this time period, most research was concerned with describing differences and similarities across space (e.g., Burgess, 1925; Hoyt, 1939).

In the mid- to late-1950s, there was a growing dissatisfaction with the descriptive nature of areal differentiation because these techniques offered no explanations for the patterns they described (Golledge, 1980). As a result, there was a shift in methodological emphasis from simply describing spatial patterns to also trying to understand and explain the processes that influence these patterns. The method of analysis that was developed shifted the focus of inquiry away from observed aggregate patterns towards the individual decisions that combine to create these patterns. The emergence of this method, which is referred to as *behaviouralism*, helped to bring research on residential mobility to the fore of geographic inquiry because, at the most basic level, residential mobility is "... a process involving [individual] household decisions ..." (Short, 1978, 420).

Over the last forty years, a considerable amount of research on the residential mobility process has been conducted. The vast majority of this research has been framed within the concepts of micro-economics and behaviouralism. These studies have attempted to explain the process of residential mobility (i.e., the circumstances or events that influence the household decision to move). The general findings of these studies are that "Households [choose to] move when they perceive their existing dwelling to be undesirable, or at least less desirable than an alternative" (Harris, 1991). There are several factors that contribute to a household's perception of its housing situation.

- *Housing Characteristics.* Several authors have discussed the impact of housing attributes on mobility (e.g., Clark et al., 1986; Michelson, 1977; Rossi, 1955). Housing characteristics that are commonly associated with residential mobility include *tenure*, *physical attributes* and *cost*. One of the most consistent findings in mobility research is the large difference in the mobility rates of *owners* and *renters*. Renters are typically found to have a much greater desire to move and are thus more mobile than owners (Rossi, 1955). Among the *physical attributes* of housing, *space* is usually found to be the dominant factor in mobility decisions. In (Goodman, 1976) crowding, which serves as a measure of space pressure, is found to influence mobility and housing quality (measured as expenditure per room) is found to have a very small impact on mobility. In (Moore, 1972) it is suggested that the impact of *housing costs* on mobility is weak and the empirical results tend to support this claim. In (Rossi, 1955) housing costs are found to be a major consideration after the mobility decision has been made and the household is considering alternative locations. If a household is faced with two or more suitable alternatives, they will select the least expensive option (Rossi, 1955).

- *Neighbourhood Characteristics.* The relationship between neighbourhood attributes and residential mobility has been the subject of several studies (e.g., Brown, 1975; Connerly, 1986; Speare, 1974; Varady, 1983). Neighbourhood attributes, which are commonly thought to influence mobility, include *quality*, *social composition*, *social ties*, *public services* and *accessibility*. Neighbourhood quality, essentially, is a composite measure of neighbourhood characteristics. While some studies (Speare, 1974; Rossi, 1955) suggest that *neighbourhood quality* impacts the mobility decision indirectly through residential satisfaction, others claim that most moves are short, within familiar territory, and reflect satisfaction with the neighbourhood (Simmons, 1968). With regard to *social composition*, while some argue that the mobility is the same, regardless of social background (Simmons, 1968), others show that the mobility process does not work in the same way for households from different ethnic backgrounds (St. John *et al.*, 1995). While some studies state that over time (i.e., as people age), people become integrated into the community in which they live through social relations and organizations, and the more integrated a person is, the less likely he/she is to break the ties by moving (Goldscheider, 1971; Speare, 1974), others show that *neighbourhood relationships* have a weak impact on mobility in comparison to factors such as tenure and age (Connerly, 1986). *Public services* include a wide range of common "goods" provided by the local government, such as police protection, garbage removal and schools. While, it is a commonly accepted belief that households move because of the poor provision of public services in some neighbourhoods (Michelson, 1977), the results of empirical studies are not conclusive (Varady, 1983). With regard to the relationship between *workplace accessibility* and residential mobility, there are reviews, which explicitly reject any relationship (Clark, 1982; Short, 1978; Simmons, 1968). However, in (Brown, 1975) the produced results demonstrated that households that did not experience a change in employment had a substantially lower annual move rate than those households that did undergo a change in job location.

- *Income.* A number of authors have analyzed the relationship between changes in income and residential movements. Generally, the results conclude that there is weak relationship between income and voluntary mobility (Speare; 1974; Pickvance, 1974; Duncan and Newman, 1976).
- *Life-Cycle.* Without debate, the life-cycle hypothesis has received the most attention of any of the theories associated with residential mobility. First analyzed by (Rossi, 1955), the life-cycle hypothesis generally states that mobility is a process by which households adjust their housing to the needs that are created by the changes in household structure that result from life-cycle transitions (e.g., the birth of a child). Empirical studies have generally provided support for the hypothesis. (Pickvance; 1974; Duncan and Newman, 1976; Clark et al., 1986; Speare and Goldscheider, 1987).

While it is possible for the mobility process to be triggered by any one of these causes, it is more likely that the decision to move is conditioned by a combination of events. The next section looks at how these factors, and different combinations of factors, have been incorporated into a modelling framework.

2.3 Modeling residential mobility

There are two basic approaches that can be used to study any phenomenon: empirical and theoretical studies. Empirical studies seek to explain processes based on observations of how they work. In studying processes of human behaviour, it is not always possible to directly observe the mechanism under investigation and, as a substitute, researchers use surveys in order to find out how individuals respond to specific stimuli. Once data for a representative sample has been collected, the researcher is able to infer generalizations of how the process operates for the entire population.

In contrast, theoretical studies attempt to explain a process by creating a simplified representation of that process (i.e., a model) and observing how the model responds to situational changes. Models can take on many forms: physical, graphical, mathematical expressions, verbal descriptions or computer programs. As with empirical analyses of human behaviour, a researcher can run into problems due to the complex nature of human decision-making. If the researcher tries to incorporate all of the variables involved in the decision making process, the model will quickly become overly complex and extremely difficult to operationalize. On the other hand, if the researcher leaves out a critical variable, the results generated could be meaningless. As a result of the difficulties in designing a representative model of human behaviour, model builders have tended to rely on the results of empirical studies in order to determine which factors are the most important in terms of explaining the process being studied. Models of the residential mobility process are a classical example of how the results of empirical studies have been incorporated into a modelling framework.

The vast majority of residential mobility models have been developed from a behavioural perspective. The behavioural method is concerned with studying human activity from the viewpoint of the individual. The primary goal of behaviouralism is to "... identify the steps involved in decisions made by individuals and [the] constraints wherein these decisions are made" (Filion and Bunting, 1991, 13). Thus, the *behavioural* method views residential mobility as a decision made by individual households. This is the approach that has been applied to models that approach mobility from an economic view as well as those that approach it from a sociological view.

Generally speaking, economic models of the residential mobility process assume that households derive utility from their housing situation. Households are assumed to consume housing so that the utility they obtain is maximized. These are standard assumptions of economic models.

The economic models of mobility that have been developed are based on the disequilibrium model of housing consumption (e.g., Goodman; 1976; Hanushek and Quigley, 1978a, 1978b). These models assume that households derive utility from the consumption of housing and all other goods (a composite commodity). Subject to a budget constraint, households will maximize their utility by consuming the appropriate amounts of these two goods. If there is a change in housing attributes or household preferences, a household will begin to deviate from its equilibrium level of housing consumption. In order to realign consumption with preferences, the household will move to a new house which satisfies better the household's needs/wants. The probability of moving is the greatest for households that deviate the most from their equilibrium consumption level and who face the lowest costs of moving. A move occurs if the utility gain from the move is greater than the disutility associated with moving (i.e., search, psychic and monetary costs).

There has been a large number of residential mobility models developed from a sociological perspective. In the sociological literature, there have essentially been two main approaches to mobility modelling: (cumulative) stress adjustment and (cumulative) inertia effects.

(Rossi, 1955) was the first to analyse the link between changes in the family life cycle and a household's propensity towards residential mobility. The empirical regularities of the family life-cycle cited in the mobility literature have been generalized into a life-cycle model of residential mobility by (Abu-Lughod and Foley, 1960). According to the life-cycle model, a household goes through a number of different stages in its lifetime. As a household progresses through these phases of development, it experiences a number of changes in its structure, which impact upon its housing needs. As a result, the household will move in order to satisfy this change in housing needs.

The stages of development in the life-cycle model are intrinsically linked to the space requirements of the household. As the household moves through its life cycle, and subsequently experiences changes in its size, it will move in order to adjust its housing situation to these space requirements.

It should be stressed that this development process is a generalization of household transition and that all families do not follow the exact same life course. Specifically, the life-cycle model generalizes the movement patterns of middle-income households. This issue has become more salient in recent years, as there has been a large increase in the number of alternative households (for example, single parents, childless couples, same-sex partners). Also, changes in a family's life cycle do not fully account for the housing considerations of the household (Kendig, 1984). Preferences alone do not dictate where a household lives and that a household's housing situation cannot be properly understood without considering their economic means and the constraints that exist in the housing market (Kendig, 1984). Also, there may be a strong correlation between changes in the family life-cycle and changes in its economic conditions (Kendig, 1984). For example, a young couple wanting to start a family may wait until such time as it can afford to do so. Therefore, while the family life-cycle model provides many useful insights into one reason behind the mobility decision, it does not fully explain the mobility process.

Another important concept in modelling residential mobility from a sociological perspective is the idea that there exist factors, which work to impede the mobility process. The concept of cumulative inertia was introduced, which is the idea that "The probability of remaining in any state ... increases as a strict monotone function of duration of prior residence in that state" (McGinnis, 1968, 716). This means that, the longer a household lives in a specific dwelling (i.e., state), the less likely it is that it will move from that dwelling.

Compared to the volume of literature which focuses specifically either on cumulative stress or cumulative inertia, there have been a negligible number of attempts to model both processes together (Huff and Clark, 1978).

Residential mobility models have evolved over time and are presently in the process of being further developed. A number of researchers have applied a modelling technique, known as microsimulation, to the study of mobility. This technique has the potential to improve significantly the modelling of the residential mobility process and is discussed in the next section.

2.4 The microsimulation technique

Until recently, urban academics and policy makers developed demographic and urban models using conventional modeling techniques. Models based on these techniques are defined at the aggregate level (macro-models), that is, they refer to groups of individuals that comprise their population. In particular, the principle feature of such a model is an n-dimensional matrix which represents the population database used by the model; the population is categorized based on a number of attributes that are of interest within the particular application and the cells of the matrix, referred to as states, contain information on the total number of individuals within each state.

It is argued that aggregate-models do not produce reliable representations of the complex behavioral relationships that household and family demographic systems incorporate; the modeling of demographic processes such as fertility, marriage and migration, generates discrepancies between projected and real world outcomes (Bartlema and Vossen, 1988). This is due to the fact that individuals are the basic decision makers with regard to these processes, not the population as a whole. Moreover, aggregate-models cannot capture the spatial representation of the population residential mobility. Consequently, the spatial representation of model results may only be made at aggregate level i.e. the housing market level. That is why, in recent years, microsimulation has been receiving ever-increasing attention as a method for modeling household decisions.

Microsimulation was first introduced by (Orcutt, 1957) and is used for the development of models at the disaggregate level, that is, it aims at the modeling of micro-units such as individuals, households and firms, and the events they encounter. In other words, in a microsimulation model each entity of the modeled application forms a record within a database and the attributes of each entity are updated based on the assumed or observed behavior of the entity.

The development of a microsimulation model involves two main phases:

1. The construction of a detailed *database* defined at the *micro-level* (i.e. consists of a list of individuals, households, etc, and their associated attributes).

2. The construction of a *mechanism* for the processing of these micro-units through a set of rules that will lead to changes in the value of the attribute associated with each individual unit.

A microsimulation model should be capable of storing the *spatial dimension of objects* (or entities, or actors) and handling the *objects' behavior over space*. In addition, it should be capable of *storing information over time* and querying this information to determine objects' future behavior and help decision-makers with future decisions.

While it might seem as though the main concern of microsimulation is with the individual actors of a given system, this approach is actually interested in determining the system's aggregate results. Whereas aggregate models make generalizations about human behaviour and model decisions at the aggregate level, the microsimulation approach models decisions at the individual level and then combines these outcomes in order to attain an aggregate result.

The microsimulation method has several advantages over the macro approach; however five of these can be considered the most important. These five benefits include: detailed descriptions of individuals, the disaggregate nature of the results, explanatory variables are subjected to few restrictions, a decreased chance of inferential error and, finally, the ease with which theories can be incorporated into the microsimulation framework.

First, the structure of the microsimulation approach allows for highly detailed descriptions of the individuals, which are part of the system being modelled. A microsimulation database is comprised of a list of individuals that are described by a number of characteristics. The macro approach, on the other hand, uses data that is in matrix form and the cells of a macro-matrix contain a count of the number of individuals that fall into specified categories. Due to data aggregation, the macro approach tends to suffer from data loss and abstraction (Orcutt *et al.*, 1968). This is not a problem in the micro case because the database contains the actual data for each individual. On a related note, the database in the micro case represents the life-histories of several individuals. Thus, the validity of the simulated events can be established by comparing the simulated and actual life-histories of individuals.

A second advantage of the microsimulation method is that the generated results are in a disaggregate form. This is advantageous for two reasons. First, results of this nature allow for flexibility in aggregation. If the researcher needs to analyse the data at various scales (e.g., local, regional, national), this is done with relative ease. Second, analysis of the data for different sub-groups within the population (e.g., sub-groups could be defined spatially or by any of the attributes in the database) allows the researcher to establish how the results are distributed amongst these groups. This is a particularly important feature in policy analysis, where it is critical to determine the gainers and losers of a certain (change in) policy.

Another advantageous feature of the microsimulation approach involves the number of explanatory variables that can be factored into the analysis. Due to the nature of the database in a microsimulation (i.e., a list containing all individuals), the size of the data structure is not a function of the number of variables, but is determined by the sample size. Therefore, a large number of explanatory variables can be factored into the analysis without any problems of storage and computation. In contrast to this, the database in a macro model, because it is in matrix form, increases in size very rapidly as more and more variables are added to the database. This means that macro-level data files can become very large with just a few explanatory variables, which can cause computational and storage problems (this problem becomes less of a factor as computer technology keeps improving).

A fourth benefit of the microsimulation method is that it involves less of a chance in making inferential errors due to heterogeneity. Aggregation tends to mask a lot of the variation that is present within data. As a result, this "hidden" heterogeneity is often mistaken for substantive relationships between variables. In microsimulation, the results are in a disaggregate form and, therefore, the variation in the results is not masked. This situation reduces the chance of drawing spurious conclusions.

A final advantage associated with the microsimulation technique is that it "... allows for the incorporation of theories that have been developed at that level as opposed to those relating to aggregates of micro units" (Clarke and Holm, 1987, 147). Behavioural theories are developed at the level of the individual and are meant to be applied to situations that focus exclusively on the decisions made by individuals. Since microsimulation models simulate the behaviour of individuals, the concepts (e.g., decision rules) developed in behavioural theories are easily incorporated into this type of framework.

Microsimulation literature includes several residential mobility models which may be grouped in two categories. The first category includes the residential location-choice models such as SwarmCity (Devisch *et al.*, 2004) and UrbanSim (Waddell, 1998). These models describe the behavior of an individual (be it a household, firm, company etc) searching for a location to settle. In these models, the demographic model is either modeled as a transition model simulating population changes using iterative proportional fitting to create households of particular types (in case of UrbanSim) or given exogenously (in case of SwarmCity). The second category includes the demographic and housing market models, which explicitly model the demographic events experienced by individuals and households, as well as their interrelationship with residential mobility. Models enlisted in this category are: HUDS (Kain and Apgar, 1985), UPDATE (Clarke, 1995), HOUSIM (Clarke *et al.*, 1989), LOCSIM (Oskamp, 1995; Oskamp, 1997), Fransson model (Fransson, 1994; Fransson and Makila, 1994), SVERIGE (Vencatasawmy *et al.,* 1999), ILUMASS land-use model (Moeckel *et al.*, 2002) and ILUTE residential mobility model (Miller *et al.,* 1987). As the importance of the demographic model was outlined in Chapter 1, within the context of this work the second category of models is of interest.

2.5 Elements of residential mobility microsimulation models

When defining a residential mobility microsimulation model, a set of issues and concepts is of importance and should be taken into consideration. The elements that such a model should incorporate are:

- *Database entities*. The *entities* that are of interest within the context of the model and thus comprise the model's database such as individuals, households, dwellings, companies, etc.
- *Space & time dimensions*. The ability of the model to store the spatial dimension of database entities, to handle their behavior over space, to store information over time and to query this information.
- *Housing, demographic and economic events.* The event attributes which are of importance in residential mobility microsimulation modeling include the occurrence of events, the sequencing of events and the event modules.
    - *Occurrence of events*. During a simulation period, each individual may experience one or more events. Various methods may be used to determine whether an individual experiences a specific event or not.

Monte Carlo sampling is such a method, and is employed by several microsimulation models. In Monte Carlo sampling (Oskamp, 1995), a random number from a uniform distribution [0,1] is drawn for each process that an individual may experience. This number is compared to the probability that the event occurs; if the random number falls in the interval between 0 and the event probability then the event occurs; if the random number falls in the interval between the event probability and 1, the event does not occur.

- *Sequencing of events.* Since each individual of the population is *sequentially* subjected to the probability of experiencing each event, the order of events is of importance mainly for two reasons (Oskamp, 1995). First the sequence may bias the number of events that occur. For example, if death is simulated before fertility, women who died may have had a child before they died. If fertility is simulated before death, births are included of women who would have died before they gave birth. Second the occurrence of one event may change the probability of occurrence of another event. For example, if marriage (union formation) is simulated first, then fertility probability increases since it depends on marital status. Most models use a *fixed order* of events. Alternatively, one strategy used in event ordering is the *random order* of events called *random dates* (Oskamp, 1995). This is a two-step process. The first step only determines if an event will occur, and if so on which of the 365 days of the year. The second step then executes the events in the order determined in the first step.
- *Event modules*. It refers to the events simulated by the model. Within a model, an individual may experience *demographic* events such as fertility, nest leaving and marriage, and *economic* events such as job change. A dwelling may experience *housing* events such as demolition, split and renovation.

- *Housing search.* The housing search attributes which are of importance in residential mobility microsimulation modeling are the dimensions of housing search taken into consideration and the housing search method used.
    - *Dimensions of housing search.* It refers to the elements of a dwelling taken into consideration during the search process. A vacancy's attributes on all dimensions of the housing search are checked against the household's aspirations, and the searching household decides whether to accept or not the dwelling. Important dimensions of housing search are *price*, *number of rooms*, *type*, *neighborhood quality*, *relative location*, *quality* and *age* of house.
    - *Housing Search Method*. It refers to the method that households use to search the housing market, and to structure their housing decisions.
- *Realization level.* There are three levels of model realization: implementation, calibration and validation.
    - *Implementation.* The process aims at generating program code in some programming language to construct a microsimulation information system that is based on the model.
    - *Calibration.* It is a process aiming at fine-tuning the model, so that it can replicate real world developments (Oskamp, 1997). It simulates the behavior of the base population living in a geographical area over a time period, for which empirical data are available. The outcome of the simulation is compared to the empirical data and the analysis of their discrepancies reveals which code modules and parameters should be fine-tuned.
    - *Validation.* The process applies the calibrated model to different data and/or geographical area and/or period, and checks whether it yields

comparable results i.e. whether the model ‘fits’ in another context (Oskamp, 1997).

The following sections discuss existing residential mobility microsimulation models and look at how the above mentioned elements have been incorporated into the models.

2.6 LOCSIM

LOCSIM (LOCal SIMulation) (Oskamp, 1995; Oskamp, 1997) is a dynamic microsimulation model that captures the developments in demographic structures and housing markets, as well as the interrelation between them. It was developed using conventional analysis and design methods, and operates at local levels and for short periods of time within a time horizon of three to five years.

Within LOCSIM, the database entities of interest are *individuals*, *households* and *dwellings*. In addition, three types of mobility by destination are distinguished: households moving within the boundaries of the housing market (*residential mobility*), households moving from the LOCSIM housing market to another one (*out-migration*), and households moving into the LOCSIM housing market from the rest of the world (*in-migration*). Most of the processes are simulated using *Monte Carlo* techniques. In some cases, the process is clearly one of choice and may therefore be modeled with a decision rules approach.

The model consists of three main sub-models (Figure 1):

- The *Demographic Development sub-model* (Figure 2), which simulates the occurrence of demographic events (e.g. fertility, marriage etc) and therefore generates housing demand. Demographic events are assigned to individuals using *Monte Carlo* sampling. The order of events (sequencing) occurring during *a one-year period* is simulated using *random dates*. The probability distributions of the various events can be age, sex or duration specific. The model sequentially exposes individuals to the risk of experiencing demographic events. As the occurrence of such an event changes the structure of the household to which the individual belongs, the household is assigned a *move tag* to denote its desire to move. The tag may be a *preferred* move tag (the desire to move is the reaction to the occurrence of the demographic event), *implied* move tag (moving is a condition for the occurrence of the demographic event) or *forced* move tag (in case of dwelling demolition, eviction, etc). All households do not have an equally strong desire to find a new dwelling; this is captured by introducing *search intensity*. Within LOCSIM the following demographic events are modeled:

- *Mortality*: at the beginning of each simulation period, all individuals are exposed to the risk of dying, based on age- and sex- specific rates of mortality. If a death occurs, the attributes of the individual and the household to which the individual belongs, are updated accordingly, and the household is assigned a move tag. If the individual is a one-person household, then the dwelling of the deceased is added to the stock of vacant dwellings.
- *Fertility*: all female members of the population that are of childbearing age (i.e. aged 18 to 45) are subjected to the probability of giving birth, based on age-specific rates of fertility. If a child is born, a new individual is added to the population database and the household is assigned a preferred move tag.
- *Nest leaving*: all individuals of age 18 and older, who are living with their parents, are subjected to an age- and sex- specific nest leaving probability. (i.e. leaving the parental home to live alone). The nest

leaver and the parental household are tagged with an implied and a preferred move tag, respectively. If the individual's housing search is unsuccessful, then the nest leaver is returned to the parental household and the move tags are removed.

- *Divorce and dehabitation*: all individuals in unions are subjected to dissolution probabilities, which are different for marriages and cohabitations and are based on the duration of the union. When a couple breaks up, the next decision is which partner moves out of the current dwelling forming a new household, and if there are children, with which parent they go. The newly formed household receives an implied move tag and the original household receives a preferred move tag. The new household returns to the original household if housing search is unsuccessful, and the move tags are removed.
- *Marriage of cohabiters*: as it is assumed that if two persons who are cohabiting marry, they will marry each other, they do not enter the partner market. All cohabiting couples are subjected to the probability of marriage. The only consequence is that marital and cohabitation status of the individuals is updated. The newly married couple is tagged with a preferred move tag but with low search intensity.
- *Partner market: marriage and cohabitation:* all individuals not living with a partner are copied from the population to a partner market file. In the partner market each individual is first subjected to a search probability to decide whether or not the individual will actively search for a partner. If search commences, it is decided whether the individual will search within the municipal boundaries (i.e. the partner market file) or outside the municipality. The current search criteria are only age and sex, but could easily be extended to include other attributes, such as ethnicity and education. If the individual's partner search succeeds, the new household is tagged

with a move tag. If the partner is from another municipality, then it is determined whether the partner moves in with the searcher or vice versa.

- *Income change:* after having been put through the demographic processes, each individual is subjected to a probability of income change. If the income of the individual changes, the household to which it belongs, is tagged with a preferred move tag.
- *Non-triggered mobility*: each household that did not receive a move tag of any kind in the previous modules is now subjected to a probability to make a move not triggered by one of the processes (demographic events and income change) simulated in LOCSIM. If a household is to move, it is tagged with a preferred move tag.
- *In-migration:* the composition of an in-migrating household is based on probability distributions by household size and age-of-the-head. When household size and age of the head are known a random household from the population with corresponding size and age is copied to create the in-migrating household.

- The *Housing Supply sub-model* which simulates exogenous changes on the supply side of the housing market. Changes in housing stock are modelled in LOCSIM using Monte Carlo sampling. Dwellings are exposed to the risk of being demolished, split, combined and renovated and having their tenure changed. The updating of housing prices (rent and market value) is modelled using exogenously induced rent and market value increase. Changes in the characteristics of the dwellings households occupy generate moving intentions.
  - *Demolition of dwellings*: each dwelling is exposed to the probability of being demolished, based on the NCBS housing statistics. If a dwelling is to be demolished, it is removed from the database and the household occupying it, is tagged with a forced-move tag, giving it a high priority in housing allocation.

- *Splitting of dwellings*: each unoccupied dwelling is exposed to the splitting probability. If splitting takes place, the database is updated.
- *Combining dwellings*: all sets of unoccupied dwellings and unoccupied neighbouring dwellings are exposed to the combine probability. If combining takes place, the database is updated.
- *Tenure change*: every rented dwelling is exposed to the probability of changing into an owner-occupied dwelling based on Dutch Housing Demand Surveys. If tenure change occurs, the resale value is determined based on the market value of the dwelling or the average market value of the dwellings in the street.
- *Renovation*: all dwellings are exposed to a renovation probability. If renovation takes place, the quality indicator of the dwelling is updated as well as the rent and/or market value, and, in case of rent-increase, the household is tagged with a preferred tag.
- *Rent and market value update*: all rents of rental dwellings in the housing database are increased by the average percentage valid for the municipality. All market values of owner-occupied housing are annually increased by 6.67 per cent.
- *Housing construction*: new housing is added by exactly specifying new dwellings in complete dwelling records which are added to the stock on January 1rst, and are directly available.

- The *Housing Market sub- model* (Figure 3), which simulates the clearing of the housing market (the match between housing demand and supply). The input of the housing search module is *stated preferences* (what households want) and the output is *revealed preferences* (what households got). Household mobility is the outcome of successful housing search. A *probabilistic heuristic search model* is used to structure decisions. The model uses Decision Plan Nets (DPNs) in conjunction with Monte Carlo techniques that parameterize the decision rules. A *decision plan* is the structured representation of a household's decision process, i.e. what the household does if the attribute of the dwelling under consideration on a housing search dimension does not comply with the household's desire. LOCSIM captures the constant and continuous allocation of dwellings to households searching the market, by executing the housing market simulation in *four cycles*. Dwellings vacated in one cycle are available to searchers in the next cycle.

- *Selecting potential movers*: all households whose structure or housing circumstances have changed and all newly formed households as well as the households they originate from are selected. Finally, all other households are subjected to a move probability to account for non-triggered mobility.
- *Selecting vacancies for internal movers*: for each dwelling, it is determined if it is to be offered to internal or to external searchers., based on the housing market competition probability by dwelling type.
- *Selecting a searching household*: within the model, households are searching the market sequentially, within an order determined by their priority in municipal housing allocation regulations. Within each group of households with the same priority, households are ordered randomly.
- *Determining the search intensity*: during a housing market cycle, each household evaluates offers until one is accepted or the number of rejects equals the *search intensity* (i.e. the maximum number of offered dwellings to be inspected by a household in one housing market cycle). The household's search intensity is estimated on the basis of its reason to move.
- *Determining the acceptance interval*: For existing households, the lower boundary of the *acceptance interval* (i.e. the range of acceptable dwellings) is the present housing situation of the searcher. For starting households, the lower boundary is set to an arbitrarily

defined minimum. The upper boundary is the aspired housing situation. Within LOCSIM, dwellings are defined by seven dimensions: *price*, *number of rooms*, *type*, *neighbourhood quality*, *relative location*, *quality* and *age* of house. On each of these dimensions an acceptance interval is estimated and in order to fall within the total acceptance interval, a dwelling has to fall in the seven partial intervals.

- *Selecting a vacancy*: the vacancy which is to be offered to the searching household is selected randomly from all vacancies the searching household is eligible for and that fall within the acceptance interval. The number of possible vacancies offered to the searching household (*arrival rate of opportunities*) is determined by municipal housing allocation regulations, the number of vacancies available, the competition of other households, and the acceptance interval of the household.
- *Determining the acceptance probability*: dwellings nearer the lower boundary of the acceptance interval have a lower probability of being accepted (*acceptance probability*) than dwellings closer to the aspirations. The acceptance probability is estimated for the selected vacancy based on the Dutch Housing Demand Survey (WBO).
- *Deciding to accept or reject*: if the household decides to reject the vacancy, it will be offered another dwelling, until either the household accepts, or the number of rejected vacancies equals the search intensity. Each rejected offer increases the acceptance probability.
- *Updating movers file and vacancy file*: unsuccessful potential movers as well as households from which a starter originated, wait for the next housing market cycle to start. Searching households that move or out-migrate, have their original dwelling vacated and are added to the list of obtainable vacancies in the next cycle.

- *Returning unsuccessful starters*: starters that fail to secure a dwelling, return to the original household, and the occurrence of the corresponding event as well as their intention to move, are cancelled. All other unsuccessful potential movers remain in the original dwelling and retain their move intention, placing them back on the market at the beginning of the next interval.

The data flow diagram technique (discussed in Chapter 3) was used for the graphical representation of LOCSIM during design. The model was *implemented* in Pascal 7.0 (a simplified version of the model was implemented due to data availability restrictions), *calibrated* and *validated*.

The model does not incorporate modules pertaining to educational and employment careers. In addition, since a market clearing mechanism is used by the model, the updating of housing prices (rent and market value) could be done endogenously.

The model is not capable of storing information over time and thus the life histories of the model's entities i.e. individuals, households and dwellings, may not be captured. Although the model incorporates the spatial dimension of population, the implementation platform chosen does not facilitate the spatial representation of model results at various levels of aggregation i.e. at residential zone and building block levels.

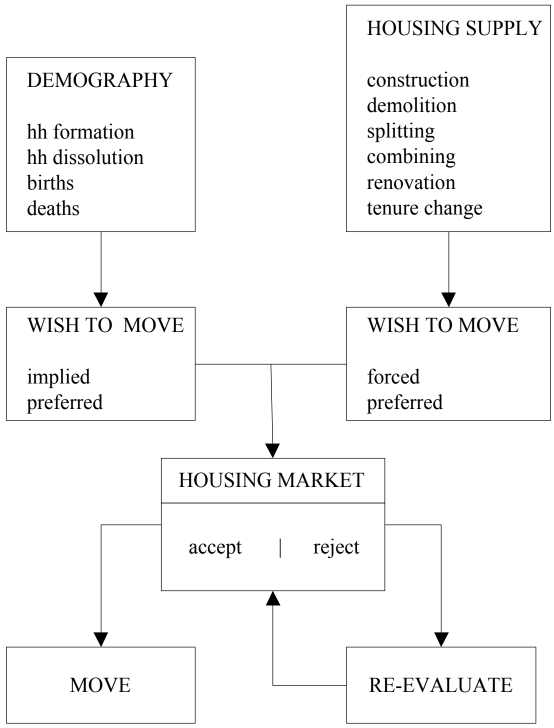


Figure 1: OUTLINE OF LOCSIM

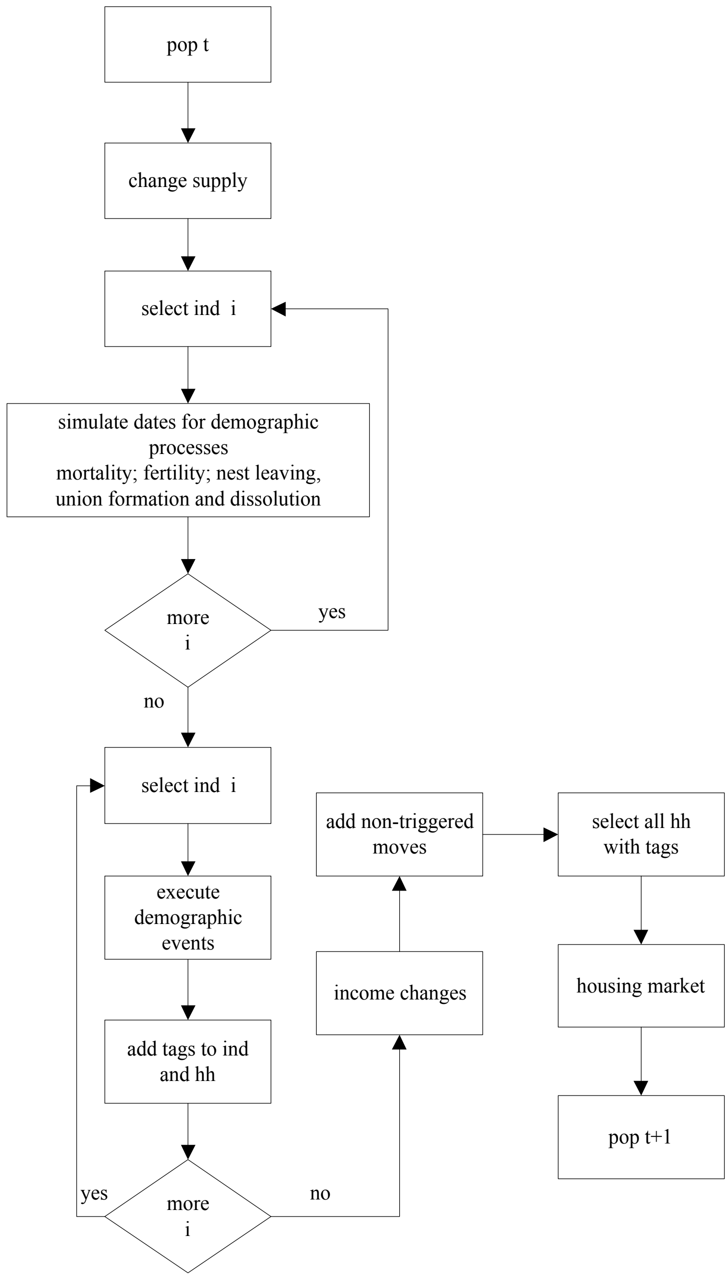


Figure 2: General outline of LOCSIM's demographic model

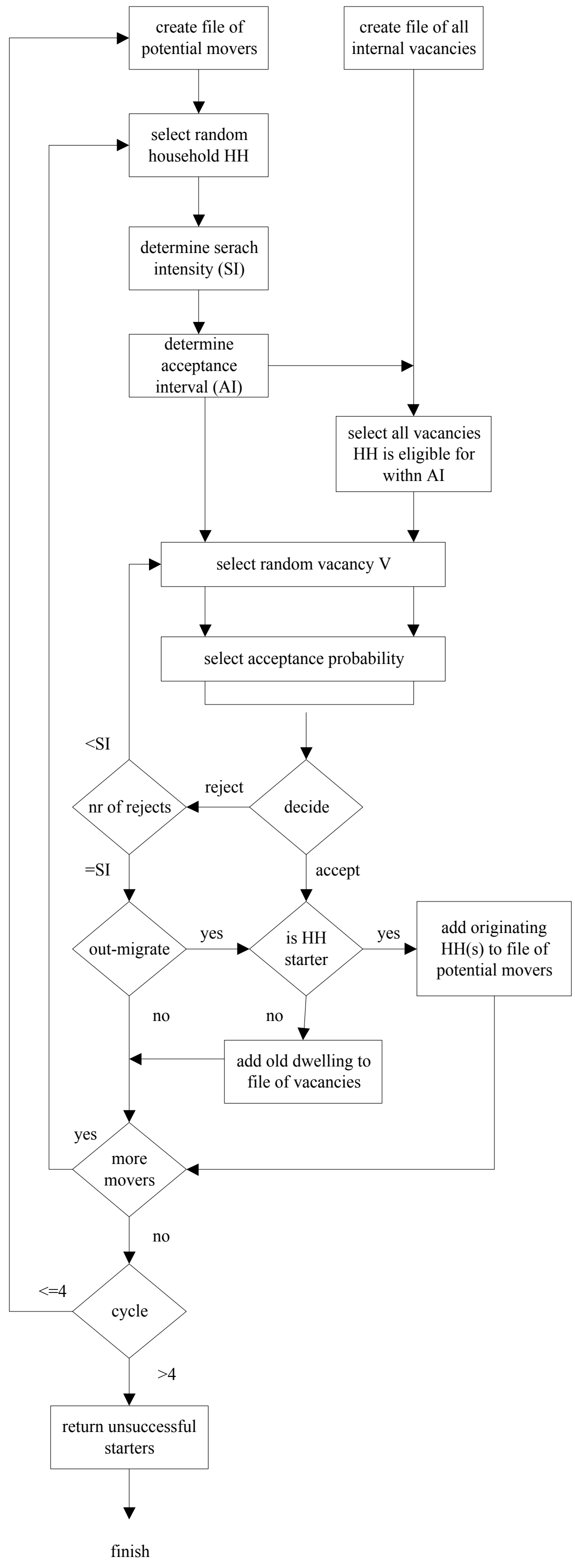


Figure 3: Schematic representation of LOCSIM's housing market module

2.7 HUDS

The HUDS (Harvard Urban Development Simulation Model) model (Kain and Apgar, 1985) simulates the processes of housing markets in large metropolitan areas, where access to workplaces has an important influence on the demand for housing. The model is designed for the Chicago metropolitan area and has been used to assess the impacts of a series of housing policy programs, such as neighbourhood improvement programs.

The model assumes that urban housing markets never achieve a long-run equilibrium but instead are constantly adjusting to new target equilibrium.

The processes modelled by HUDS are simulated at the *dwelling* unit and *household* levels. The database used contains a permanent record of each dwelling unit in the Chicago sample. If a unit is occupied, the dwelling unit record also includes attributes which provide a description of the occupant(s).

Three distinct geographic specifications are employed: the *residence zones* are aggregates of census tracts and are contained within *residential districts* which are defined on the basis of neighbourhood quality; the residential districts are further aggregated in order to specify *workplace zones*.

Dwelling units are classified as one of 50 housing bundle types, defined by 10 structure types and 5 levels of neighbourhood quality, each housing bundle produces a quantity of structure services. They are also described in terms of quantities of maintenance capital and structure capital.

The model consists of three main sub-models (Figure 4):

- The *Demand Sector* is simulated by eight sub-models that capture changes in household demographics, employment, the migration preferences of households, new household formation and the demand for specific housing bundles and tenure choices.

- The *Demographic sub-model* simulates the aging of the household head, changes in household size and changes in household income for each household. The age of the household head is incremented by one at the beginning of each year. Changes in household size are determined in a probabilistic manner given the household's size and the age of the household head. Changes in income are determined in a probabilistic manner given the household's previous income, as well as the race and age of the household head.
- The *Job-change sub-model* uses exogenous estimates of changes in the level and composition of employment in the 20 workplace zones to determine if the primary worker in a household will remain at his current job, find a new job, become unemployed or leave the labor force entirely. When the primary worker is simulated to find a new job, the sub-model determines the industry, occupation and workplace zone of this new position.
- The *Movers sub-model* selects those households that will vacate their dwelling in the current simulation period and participate in the housing market. The selection is determined by conditional probabilities given the age of the household head, tenure of the household, change in employment and change in household size. Households that are selected as movers are placed on a separate demand list along with new households and in-migrants.
- The *New household sub-model* simulates the in-migration of households to the region and adds new households to meet exogenous projections of aggregate population characteristics. The primary worker of each household is assigned to an industry, occupation and workplace zone (job change sub-model).

- The *Demand sub-model* assigns each household on the demand list to one of the 50 housing bundle types. The assignment of specific households to a particular housing bundle are determined probabilistically given income, life cycle, race, as well as the workplace location of the primary worker, the gross price of each housing bundle and the number of available units in each housing bundle.
- The *Tenure sub-model* determines probabilistically which movers will own or rent their new dwelling given past tenure, household income, the age and race of the household head and the type of housing bundle previously consumed.

- The *Supply Sector* is simulated by five interrelated sub-models which capture the decisions of landowners, homeowners and developers. Housing services are viewed as a bundle of attributes, which include neighborhood quality, accessibility, structure type and the quantity of structure services provided. The model specifies an index of neighborhood quality, which is a function of the average quantity of structure services provided by the dwelling units within each residence zone.
  - The *Expectations sub-model* estimates the decisions to make highly durable capital investments, which depend heavily upon how investors envision the future. These decisions are modeled as a function of projections of future rents. The model also estimates aggregate levels of target demand.
  - The *New construction sub-model* estimates the profit levels associated with building each type of structure in each residence zone, ranks these construction options based on the profitability measures and then allocates the construction activities in the order of decreasing profitability until the demand targets have been met.

- The *Structure conversion sub-model* determines which of the already existing structures will be replaced by the construction of new units. Due to the fact that these type of changes are quite costly, they represent the least common type of supply activity. The model assumes that projects of this nature take a minimum of two years to complete.
- The *Structure services sub-model* simulates the decisions of property owners with regard to capital outlays, which are used to maintain or increase the level of structure services associated with each building. It is assumed that a minimum level of maintenance capital is required every time period in order to preserve the level of structure services associated with a given building.
- The *Capital improvement sub-model* simulates the decline in value of all structures and the subsequent improvements made to them. The model also determines the level of maintenance capital required to produce the quantity of structure services calculated in the structure services sub-model. Finally, the model simulates demolition and abandonment decisions when, respectively, structure capital falls below 50% of the quantity included in a new structure and when the rental income of a unit no longer covers the necessary operating costs.

- In the *Market Sector* the demand and supply estimates calculated in the previous models are brought together in order to assign each household on the demand list to a residence zone. The model treats each of the 50 housing bundles as a discrete sub-market. It consists of three sub-models.
  - The *Market sub-model* uses the housing bundle assignments to simulate spatial competition among households by solving a linear programming problem which minimizes the total accessibility cost for each workplace. Households are matched to dwellings in such a

way that the total accessibility costs are minimized for each housing sub-market.

- The *Structure rent sub-model* calculates current periods rents for each housing bundle type in each residence zone and land values for each residence zone.
- The *Structure services rent sub-model* determines the quantity of structure services supplied by each structure during the current simulation period and the market rents paid for them.

No graphical representation of the system's design model exists and thus the model defined, as well as the interaction among its sub-models, is not easily discerned. The model was *implemented, calibrated* and *validated*.

The major weakness of the model is that it does not model the decision-based demographic processes such as fertility, union formation and union dissolution; Changes in household size are determined in a probabilistic manner given the household's size and the age of the household head.

Moreover, the model is not capable of storing information over time and thus the life histories of the model's entities i.e. individuals, households and dwellings, may not be captured. Although the model incorporates the spatial dimension of population, the system does not provide spatial representation of model results at various levels of aggregation i.e. at residential zone and building block levels.

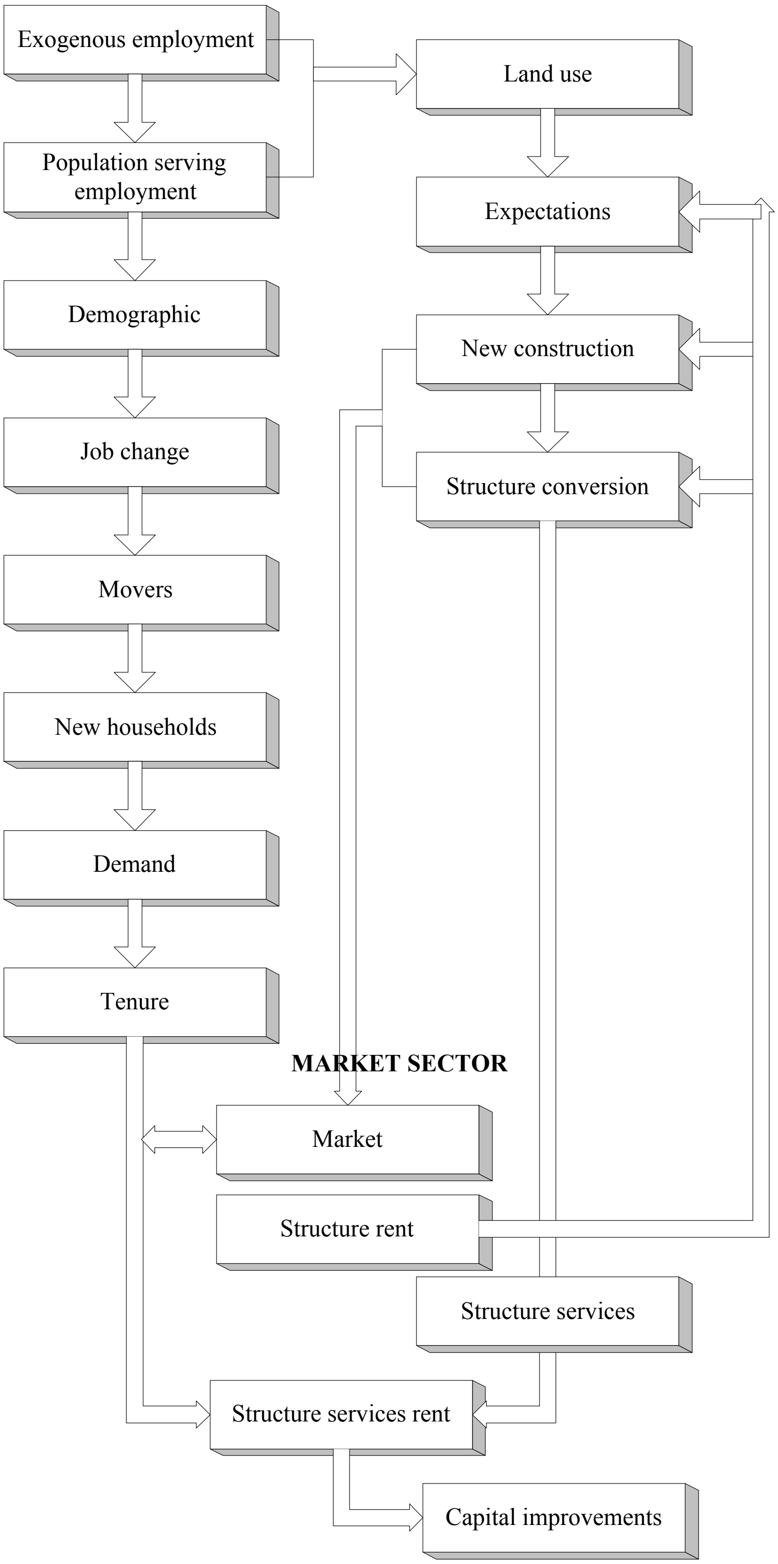


Figure 4: Sequence of major sub-models included in the HUDS model

2.8 Fransson model

The microsimulation model defined by (Fransson, 1994; Fransson and Makila, 1994) is *time geographically dynamic* and connects changes in the macro world to individual behavior as well as individual preferences, resources and decisions to the restrictions and opportunities in the market.

The model aims for a synthesis of what is already known about the different processes in housing, based on empirical studies as well as theoretical arguments. It can be used as a tool for experimenting with economic and political changes towards different groups in society as well as for analyzing the consequences of such changes.

The most important terms are *population*, *household*, *housing* and *new construction*, the *match* between *supply* and *demand* as well as *income* and *housing expenses*. The households match their demands with the supply of the market (Figure 5).

Within the model, the area of interest is divided into statistical areas that represent the geographical dimension in the model (each dwelling is linked to a particular statistical area) and the actions of the individuals on the local housing market of interest during a time period are modeled.

The behaviour of the *individual* is modelled by probability distributions and by deterministic and stochastic rules. It is assumed that social actions, together with demographic and economic events, only occur for each individual once a year.

The model consists of a number of co-operating modules (Figure 6):

- *Housing module*: housing expenditure on the different types of tenure are given exogenously (sample surveys) to each single dwelling in the model depending on period of construction and housing size. The change in number and type of dwelling is given exogenously for each year simulated. Each vacant dwelling is marked in a "pool" according to the dimensions size and expenditure.

- *Population (Demographic) module*: demographic changes such as the *death* or the *birth* of an individual as well as *in-migration* and *out-migration* affect the demand for housing. Death, birth and out-migration rates are represented by known probability distributions. The process of in-migration is represented by exogenously given empirical data as in known probability distributions and given numbers.
- *Household module*: consists of the processes that can alter the number of members of a household.
  - *Nest leaving among the young*: the opportunities of the young to leave the parental home are, among other things, dependent on the housing market supply and the resources available. Within the model, the intention of young individuals to move is modeled separately from the actual move.
  - *Divorce decisions*: the separation process is divided in two. In the first part all couples are considered to have an age specific risk of separation, which may result to a decision to separate. In the second part, the separation takes place, i.e. at least one of the partners finds a new dwelling.
  - *Moving together*: the process comprises four steps, that is, look for a partner, find a partner, accept a partner, move in with this partner to form a household. An individual can meet potential partners at work, at the place of a friend or relative or at a more open place such as in pubs or at dance locations. A man and a woman can live with their parents but still be considered as a couple.

- *Changing demands module*: a household may look for new housing due to social aspirations, to move in a more attractive neighborhood or aging. Households on low housing expenditure rate are assumed to request housing in more attractive areas for socio-economic upgrading. Households, whose current dwelling expenses exceed the household's housing budget, request a less expensive dwelling.
- *Income and educational module*: the income of the individuals is changed according to educational level, part-time work, unemployment, income career, pension and housing allowance. The educational level constitutes the base for the income career realized, which is in turn expressed as changes to the individual's salary over time. Changes to the income are affected by the general income development nationally.
- *Search and decision module*: households use a hierarchical strategy. First the household decides what housing area or part of the city is preferable for new housing and then maps the vacant dwellings. The requested dwelling has to fulfill certain criteria, which are matched against the structure of vacancies on the local housing market at the time of housing search. The household makes its decision based on only the most important criteria and on a rather limited number of vacant dwellings. The *rule of domination* is used to structure decision making of the households, that is, the dwelling attributes are compared across all the options, in that one attribute at a time is compared on all dwellings. When the household finds a new dwelling, its current one is transferred to the supply of vacant dwellings. As the structure of vacant dwellings varies during a year, households are allowed to search several times within the same year.

The model has been applied to the municipality of Ga vle, 200km. North Stockholm, Sweden, for the time period of 15 years (1976-1990). Data on individuals and dwellings according to the National Census in 1975 constitute the initial population of the model. In addition, national data as well as data for the municipality under study are used in the model. The results of the model were compared to census data on an aggregated level (*calibration- validation*).

No formally defined technique was used for the graphical representation of the system's design model, and thus several issues concerning the modeled processes, as well as the interaction among the sub-models, are not easily discerned.

Although the model incorporates the spatial dimension of population, the system does not provide spatial representation of model results at various levels of aggregation i.e. at residential zone and building block levels.

Finally, since a market clearing mechanism is used by the model, the updating of housing prices (rent and market value) could be done endogenously.

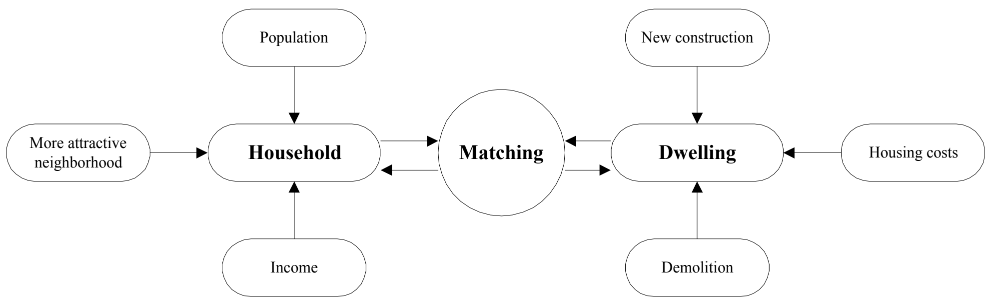


Figure 5: The main components deciding the moves of households in Fransson model

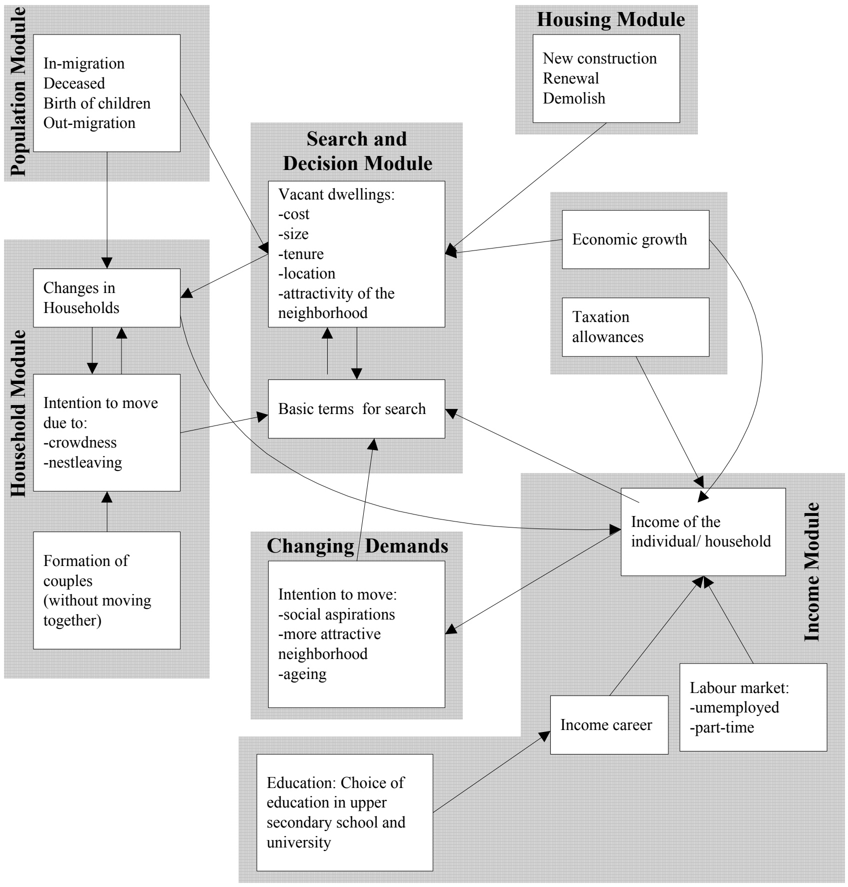


Figure 6: Relations between the modules of the Fransson model.

2.9 SVERIGE

SVERIGE (System for Visualizing Economic and Regional influences in Governing the Environment) (Vencatasawmy *et al.*, 1999) is a *dynamic spatial* microsimulation model for Sweden based on households. It generates events for individuals through the interplay of deterministic models of individual behavior and a Monte Carlo simulation. The behaviors are functions of individual, household and regional socio-economic characteristics usually included as independent variables in discrete choice models or simply as categories used to estimate transition matrices that describe the probability of moving from one state to another. The model is dynamic as the evolution and development of the individuals occurs in chronological order, with initial conditions being changed for subsequent periods by counters and simulation. In addition, the model is spatial, and thus, not only makes life-cycle transitions depended on a spatial context but also models individual spatial transitions such as internal migration.

The SVERIGE model is shown in Figure 7. Grey shading indicates that the module is operational in SVERIGE v. 1.0. Therefore, the current version of the model contains 10 modules. Each module consists of a series of discrete or continuous variable equations, transition matrices, or rules that determine the occurrence of specific events in a person's life. When individual events occur, personal attributes are updated. In this version of SVERIGE, all the events occur at yearly intervals.

The modules of the model are:

- *Mortality module.* It is used to terminate lives in the model. There are two sets of mortality equations. The first set is for individuals under the age of 25 years and is estimated as an exponential time trend from historical mortality rates grouped by age and sex. The second set is for individuals between the ages of 25 and 127 years. These equations are logistic with socio-economic attributes used as independent variables. If a person is 127 years or older, his/her life is terminated. In addition to removing individuals from the simulation, deaths directly trigger both personal and household changes. For instance, when a death occurs, the marital status of the surviving spouse changes from "married" or "sambo" to "widow." In addition, if the deceased is a single or parents die, then the under-aged children (less than 15 years) are assigned to an orphan queue and allocated to a family in the same labor market region at the end of the year. Children older than 15 years old automatically become a head of household and leave home. Attributes such as education level, employment status, and marital status influence an individual's probability of dying in any year.

- *Fertility module.* The primary role of the fertility module is to create new lives for simulation in the microsimulation model. Upon birth, each infant is assigned a sex based on a fixed probability of 0.5134 of being a male. Only woman aged 15 to 44 years old are at risk of giving birth. Equations are estimated for two groups: married and unmarried woman. The group "unmarried" includes singles, sambos, widows, and divorcees. All births produce a single baby. Fertility behaviors in SVERIGE are influenced directly by individual and household attributes such as employment, marital status, age, family earnings, and education.
- *Education module.* The model uses a series of logistic regression equations to determine completion of Compulsory school, completion of Gymnasium, entry to Komvux (an adult education program), persistence through Komvux, entry to College, persistence through College, entry to Graduate school and persistence through Graduate school.
- *Employment and earnings module*. The module estimates the amount of time each individual between the ages of 16 and 65 is employed during the year and his/her wage rate. It consists currently of four sub-modules. Initially, a logistic regression equation determines the likelihood that a given individual is employed at all during the year. For those who are simulated as being employed, the next two sub-modules determine the amount of weeks worked utilizing a logistic regression equation to determine full-time workers and a transition matrix to determine the number of weeks worked by part-time workers. The final module estimates the average relative wage rate (the ratio of wage rate to average wage rate) for each employed individual. Annual earnings are computed as the product of weeks worked and wage rate.

- *Cohabitation and marriage module.* The marriage module creates common-law marriage (sambo) for selected unmarried individuals over the age fifteen and Christian marriage partners for sambo couples. The module consists of three sub-modules. The first sub-module (sambo decision) determines whether a person is eligible for sambo or not. The second sub-module (cohabitation compatibility) computes an index of compatibility for pairs of eligible singles based on their attributes and matches pairs on the basis of the compatibility index using a heuristic matching algorithm. Each woman is assigned an index based on her age, education level and country of birth. The indexes are then sorted from highest to lowest. A man is then drawn at random from the pool and an index is computed using lookup tables. The two indexes are compared and the nearest available woman is found. The final sub-module (marriage decision) determines whether cohabiting individuals will get married. In Sweden it is very uncommon for people who move together to marry immediately soon after. Marriage usually happens after a long period of cohabitation. Couples who marry directly make up less than 5% of the total couples moving together in a given year. Because of this fact, only cohabiting couples are eligible to marry in the model. Moreover, cohabitation triggers the movement of female partners to the male partner's home.
- *Leaving home module.* The module determines whether a person will leave the parental home to start a new household. Children are eligible to leave home, beginning with the age of 16. If someone has still not left home, then he or she is forced to do so at the age of 30.
- *Divorce module.* The module dissolves sambo and marital relationships. Seven groups are used: couples without children (age groups 15-20, 21-30, 31-127), couples with children (age groups 15-20, 21-30, 31-127) and couples in which one spouse immigrated in 1990 (age group 15-

127), where the age group refers to age of the female spouse. Divorce results in persons being assigned new marital statuses (from "married" to "divorced" or from "sambo" to "single") and makes them eligible for remarriage. Also, it triggers a number of other microsimulation events, including movement of the former husband out of the marital dwelling, re-allocation of children to each new household, and de-coupling of household earnings. Currently, children are assigned to the female partner.

- *In-migration module.* Given an initial stock of immigrants, look-up tables of historical probabilities are used to assign the age, sex and marital status of the head of the household. Given these three characteristics, a comparable individual from a pool of immigrants is chosen to clone this new immigrant family. The economic and family characteristics of this clone are thus assigned to the new immigrant household head. The labor market region where the immigrant will settle is then estimated using a transition matrix based on immigrant settlement characteristics for the different immigrant groups. The family is then placed into a queue of families looking for residential locations (100 meter square land tracts). They are assigned a particular location in the *migration module*.
- *Emigration module.* The module determines who will leave Sweden as emigrants. Statistics Sweden defines an emigrant as anyone who intends to settle abroad for at least one year. Four groups were used, based on this empirical analysis: everyone between the ages of 16 and 21 years; single individuals between the ages of 22 and 58 years; couples for which the head is of age between 22 and 58 years, and everyone of age greater than 59 years. The unit of observation in the third group was the head of household. In the remaining groups the unit of observation was an individual.

- *Internal migration module*. This migration module tracks the movement of households and individuals to an accuracy of 100 meters. Movement can be either *inter-regional* or *intra-regional*. Sweden can be divided into 108 labor market regions where commuting between labor markets is minimized while commuting within labor market region is maximized. A move is called intra-regional when it happens within a labor market and is called inter-regional when it happens between labor markets. In this version emphasis has been given to inter-regional migration. Intra-migration is modeled by heuristic rules. The inter-regional migration decision is divided into three stages: Decision to move, choosing a labor market, and choosing a 100-metre square. In discriminating between movers and non-movers in stage one, both personal and regional attributes are used. In the second stage only regional attributes are used and in the third stage an attraction measure based on personal characteristics is used. In intra-regional migration, the first two steps are ignored.

A diagram technique was used for the graphical representation of SVERIGE during design. The model was *implemented*.

The major weakness of the model is that it does not model dwelling search: it is included in future work. Moreover, the model does not incorporate the housing supply of the housing market under examination.

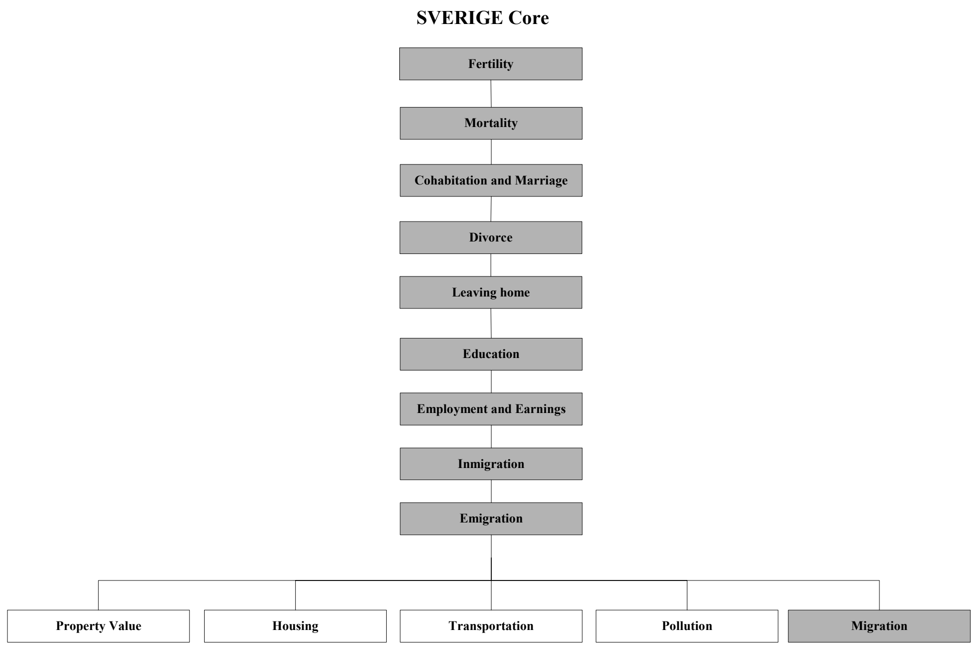


Figure 7: The SVERIGE model

2.10 UPDATE

A demographic microsimulation model (Clarke, 1995) used to revise the attributes of a population of individuals and households over time at the small area level. The purpose of the model is to provide demographic profiles in non-census years so that planners can properly allocate resources.

The model uses a synthetically generated population of *individual* and *households*. The population of each postal sector is processed sequentially through a series of demographic processes. Each postal sector is treated as an isolated system receiving in-migrants and releasing out-migrants. The number of in-migrants is dependent upon events within the specific postal sector only, and is not contingent upon any events that occur outside of the given region. The solution procedure used by the model involves *list processing* and *Monte Carlo sampling*.

The following processes are simulated:

- *Mortality*, which is conditional upon age and sex. Mortality probabilities are estimated based on national death rates.
- *Maternity*. Maternity probability is estimated by exposing women to the risk of giving birth in each year. Births are determined by sampling from a probability distribution of the number of births per maternity. They are concentrated at one birth per maternity, but also take into account the chance of multiple births (e.g. twins). New born infants are assigned a sex based on local probabilities of a male or female. Births depend upon age and sex, but also upon marital status and ethnicity.
- *Marriage/divorce*. The marriage process involves the matching of two individuals with appropriate characteristics. The model creates two pools of individuals (male and female pools) which are sampled from all of the individuals deemed to be eligible for marriage. The probability of entering one of these pools is dependent upon age, sex and marital status. These two lists are then aggregated into appropriate age groups and an allocation procedure is carried out. The probability of divorce is defined as a simple function of age and sex of the household head. In the event of a divorce, it is assumed that the husband initiates a search for new housing and both former partners become eligible for remarriage.
- *Migration*. Three types of migration processes are addressed: migration of wholly moving households, migration of independently moving individuals and the probability of intra- or inter- regional move. Migration probabilities for households in the base year, classified by household type and age of household head, are estimated for the district containing the postal sector of interest. These estimates are adjusted to reflect mobility patterns in the postal sector of interest, changes in mobility between the base and current years, and the chance of being forced to move due to demolition. For households that do not move as

an entire unit, each eligible individual is exposed to the risk of moving as an individual or as part of a subset of household members. Thus, the model creates a pool of moving entire and part households that will search for vacancies on the housing market. Movers are allocated a destination inside or outside of their origin postal sector by sampling the appropriate probabilities of staying or leaving. Next, they are matched with a vacant dwelling in the zone of destination based on a probabilistic matching of housing characteristics to housing needs.

The major weakness of the model is that it does not model the housing supply of the housing market under examination; no housing supply module is included in the system.

2.11 HOUSIM

A dynamic microsimulation model (Clarke *et al.*, 1989), that explores the processes of asset accumulation, mobility and subsidy effects in the owner-occupied sector of the British housing system.

A detailed database of *households* incorporates a large number of demographic, economic and housing attributes specified from the 1976 English House Condition Survey (EHCS) such as tenure, age, sex and marital status of household head, and house purchase price. The model updates or simulates changes in these attributes on an annual basis for the ten year period between 1976 and 1986.

HOUSIM makes use of *list processing*, whereby each household in the database is tested for a series of potential events and any changes are updated accordingly. Transitions, such as residential mobility, are determined via *Monte Carlo sampling*, which employs the use of transition probabilities stored in a series of look-up tables.

The model consists of a number of steps that are repeated for each year of the simulation:

- *Annual Update of Housing and Socio-Economic Attributes.* For each year in the simulation period, the following attributes are updated: the *age* of the household head is incremented by one at the beginning of each year; the *income* of the household head is assumed to increase by average regional rate of wage inflation across all occupations; the value of any given house; the amount of annual mortgage repayments; the amount of the mortgage outstanding; the accumulation of assets and capital; the tax relief that owner-occupiers with a mortgage are entitled on the associated interest payments; the household capital gains exemption.
- *Mobility Sub-model*: simulates the movement of households between dwelling units and tenures. The model distinguishes between two types of households: those that are already in owner-occupation and those in "other" tenure categories. Households already in owner-occupation are assigned probabilities of remaining owners, while households that are not owners are assigned probabilities of becoming owner-occupiers. For households simulated to move, the following mechanism is used to determine the characteristics of the new dwelling: first, the house price is calculated for households that are already owner-occupiers and for first time buyers; next, the mortgage amount is calculated; finally, the term of the mortgage is determined.

The major weakness of the model is that it does not model the decision-based demographic processes such as fertility, union formation and union dissolution. Therefore, residential mobility is determined via *Monte Carlo sampling*, which employs the use of transition probabilities stored in a series of look-up tables. Moreover, the model does not incorporate the housing supply of the housing market under examination.

2.12 ILUTE residential mobility model

A prototype microsimulation model (Miller *et al.,* 1987) of urban residential mobility defined as part of the Integrated Land Use, Transportation, Environment (ILUTE) modeling system (Miller *et al.*, 2004).

It is a typical *time-driven* stochastic simulation model, which simulates the evolution over time of the *spatial* distributions of *households*, *building stock*, and *place of residence – place of employment linkages*. The model views urban systems as being inherently in a dynamic state of incremental adjustment to a range of internal and external forces, rather than on any explicit concept of equilibrium.

The model has been defined for application to the Toronto Census Metropolitan Area (CMA). The study area has been divided into 133 zones which are combinations of census tracts.

Within the model, residential mobility is viewed as resulting from a divergence (residential stress) between a household's residential needs and wants and the satisfaction of these needs and wants which the household receives from its current residence. If this residential stress is great enough and if the housing market appears to provide opportunities for relieving this stress, then the household is likely to enter the housing market in search of a new residence, and, if successful in this search, actually move to a new location.

Household residential needs and wants are assumed to depend upon household demographic characteristics, household participation in the labor force, characteristics of the household's current dwelling unit, as well as perceived characteristics of alternative dwelling units, and characteristics of the household's current neighborhoods, as well as perceived characteristics of alternative neighborhoods.

The basic time increment used in the model is one year. The initial system state for the model is defined by the 1971 Census. Within the basic time increment of the year, residential mobility decisions are modeled on a monthly basis.

The model's database consists of a number of *two- and three- dimensional matrices* which collectively define the system state at each point in time. Two types of data are maintained: distributional information i.e. matrices that define spatial distributions of persons, workers, households and "linkages" between them, and activity rates i.e. matrices that define marriage, mortality etc rates. Individual households are *synthesized* as required from these aggregate distributions. A *household* is specified by its household type, zone of residence, dwelling unit type, current dwelling unit market price, the number of persons and workers in the household, the occupation and place of employment of each worker, and the household income. Every demographic and employment-related change generated in a given time period in the model requires the synthesis of "the" household associated with this change.

The model consists of four major components: life cycle (demographic) changes, labor market (employment) changes, housing stock updates, and residential search and migration.

- *Life cycle components*. The following demographic processes are likely to influence residential mobility in a significant way:
  - *Marriage*: the number of marriages to occur in a population of eligible persons (i.e. single, divorced, widowed) can be determined as a Binomial process given the average marriage rates. Information associated with each marriage is synthesized and updated, and the corresponding households are added to the "potential movers" file for later evaluation by the residential mobility sub-model.
  - *Births*: they are obtained by sampling from a Binomial distribution given fertility rates by age group. Information associated with each birth is synthesized and updated, and the corresponding household is added to the "potential movers" file.
  - *Household dissolution*: they are generated by sampling from a Binomial distribution given the number of households in each household type and the probability by household type of an individual household dissolving. Information associated with each household dissolution is synthesized and updated, and the resulting households ("leavers" and "stayers") are added to the "potential movers" file.
  - *Deaths*: a random number of deaths is generated from a Binomial distribution given mortality rates and population groups by sex and age. Information associated with each death (household, worker, dwelling etc) is synthesized and updated.
  - *Aging*: age categories correspond to each year of life with only exception the last category "75 years of age or older".
  - *In-migration*: the total number of in-migrants to the study area each year is an exogenous input to the model.
  - *Out-migration*: the current version of the model does not generate any out-migration of households from the study area.

- *Labor market components*. Employment changes within a household have the potential to influence residential mobility. Changes in the market demand for workers are currently determined by exogenous inputs into the model. Labor force reactions to such changes are endogenously generated.
  - *Job changes*: exogenous layoffs generate additional unemployed workers each year by place of work and occupation group. Fired workers are estimated according to rates by occupation group. Some employed workers will change jobs in a given year, where the number of such job transitions is randomly generated based on average transition rates by occupation group. Each worker changing jobs is permitted a choice of a new occupation group from among feasible alternatives given the current occupation, weighted by the number of vacancies. The new place of work is then selected. The household to which each transitioning, layed off or fired worker belongs, is then stored in the "potential movers" file.
  - *Entry and exit*: the number of workers entering or leaving the labor force each year can be explicitly estimated using census labor force participation rates by sex and age. Households experiencing these changes are stored in the "potential movers" file. New workers, currently unemployed workers and new in-migrants compete for the job vacancies available within each occupation group.
- *Housing stock components*. Important "actors" in the housing market include "developers", "builders", "landlords" etc whose decisions determine the levels of new building stock construction, as well as the extent of old building stock conversions, renovations and demolitions occurring in each time period. The current version of the model, however does not include any endogenous component of the supply side of the housing market. Rather, changes in the supply side of the housing

market are introduced into the model as exogenous inputs during any month within the year being simulated.

- *Residential search and migration.* The residential mobility sub-model is based on the concept of *continuous search*, i.e. every household in the urban area is always active in the housing market in that it is constantly accumulating at least some information about the housing market. Each household accumulates information at its own level of activity, called *search intensity.* A household is said to be "actively" searching when this search intensity is high and "passively" searching when it is low. When the "Anticipated Residential Utility" (ARU) of the hypothetical alternative residence exceeds the "Present Residential Utility" (PRU) of the currently occupied residence by a sufficient amount, then it is assumed that the household will transition sharply from passive search to active search. The *cusp catastrophe surface* of the catastrophe theory is used to model continuous residential search. All households not already in the "potential movers" file which fall into the active search range are added to the file. Each household in the file is processed sequentially. For each household an initial search area is defined. Next, ARU and PRU are estimated and the catastrophe model is then used to ascertain the search intensity. Households with active search intensity are the only ones retained in the file as active searchers. During the current month, an active searcher will consider either five or ten new vacancies, which are selected at random from the set of vacancies available within the household's search area constrained by the household's financial capabilities. For all those vacancies offering more residential utility than the current dwelling, a bid is formulated. The best alternative to bid on is the one that maximizes the sum of residential utility and the utility associated with all other goods which could be purchased if this bid was successful. During the market clearing process, records are selected at random from the file of bidding households for the current month. If a bid is accepted by the seller, then the household is successful in its residential search. In this case the household is removed from the "potential movers" file and the corresponding information is updated.

The major weakness of the model is that it has *not* yet been *implemented* and tested with real data.

2.13 ILUMASS land-use model

A model defined as part of ILUMASS (Integrated Land-Use Modelling and Transportation System Simulation) The land-use model of ILUMASS will be based on the land-use parts of an existing urban simulation model, but is to be microscopic like the transport parts of ILUMASS. The model works with synthetic micro data which can be retrieved from generally accessible public data.

The model consists of the following modules:

- *Household Formation.* The module models the evolution of household attributes associating each household with a particular life (Figure 8). In the household formation module the following household events are modeled simultaneously for households and household members: *birth, ageing, death; new household, dissolution of household; marriage/divorce, cohabitation/separation, separation of child, person joins household; new job, retirement, unemployment; change of income.* Even though household formation events in reality are the outcome of more or less rational decisions, most of them will not be modeled as decisions but simply as the result of the passage of time, i.e. as transitions. Typical transitions are changes of the state of a household with respect to age or size conditional on the relevant probabilities for events such as ageing/death, birth of child, relative joins or leaves household. Also clearly choice-based events such as marriage or divorce are modeled as transitions because the causal chain behind them is not represented in the model. Some events result in the dissolution of households or the creation of new households. Other events, such as a new job or un-employment are triggered by external events such as hiring or firing in the labor market represented in another part of the model not described here. Change of income is a consequence of employment-related events. Beyond these straightforward relationships there is wide scope in the model for introducing more complex interdependencies between household and economic events. For instance, the rise of dual-worker households may be in part a life style choice and in part a necessity dictated by rising housing costs and stagnant real incomes. Children may delay new household formation or marriage. Childbearing may be postponed based on some combination of life style preferences and response to housing cost and income expectations. The role of labor market expectations in shaping these choices is an area of considerable policy implication.

- *Housing Choice.* The module models location and housing choice decisions of households who move into the region (*in-migration*), move out of the region (*out-migration*), move into a dwelling for the first time (*starter households*), or have a dwelling and move into another dwelling (*moves*). Dwellings are affected by ageing and by decisions on new construction, upgrading and demolition modeled in other modules not described in this context. The housing choice model is a Monte Carlo microsimulation of transactions in the housing market. A market transaction is any successfully completed operation by which a household moves into or out of a dwelling or both. There are two types of actors in the housing market (Figure 9): households looking for a dwelling ('dwelling wanted') and landlords looking for tenants or buyers

> ('dwelling for rent or sale'). The household looking for a dwelling behaves as a satisfier, i.e. it accepts a dwelling if this will improve its housing situation by a considerable margin. Otherwise, it enters another search phase, but after a number of unsuccessful attempts it abandons the idea of a move. The amount of improvement necessary to make a household move is assumed to depend on its prior search experience, i.e. to go up with each successful and down with each unsuccessful search. In other words, households adapt their aspiration levels to supply conditions on the market. The attractiveness of a dwelling for a household is a weighted aggregate of the attractiveness of its location, its quality and its rent or price in relation to the household's housing budget. The attractiveness of the location and the quality of the dwelling are themselves multiattribute encompassing relevant attributes of the neighbourhood and of the dwelling.

The major weakness of the model is that it has *not* yet been *implemented* and tested with real data.

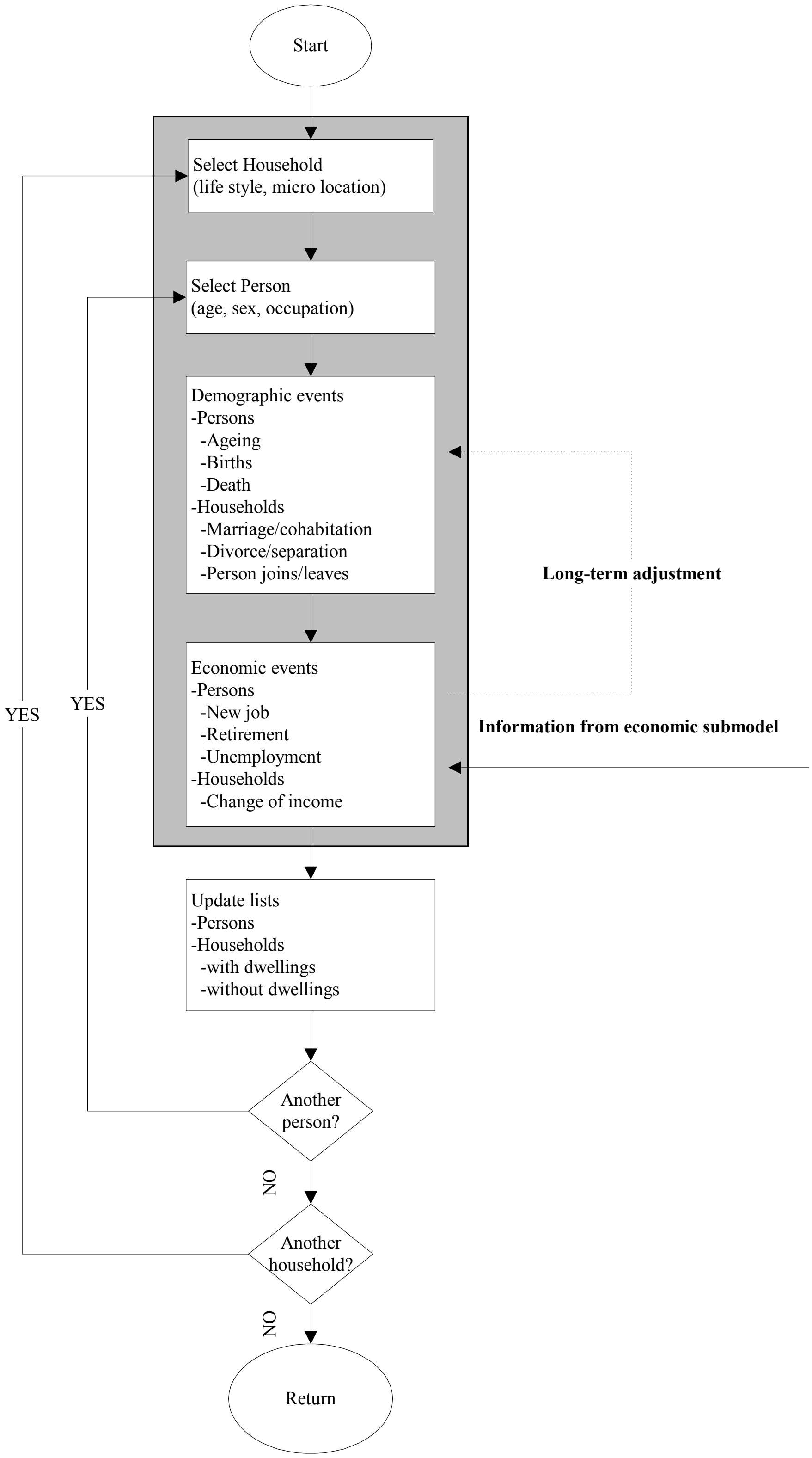


Figure 8: Microsimulatiion of household formation – ILUMASS

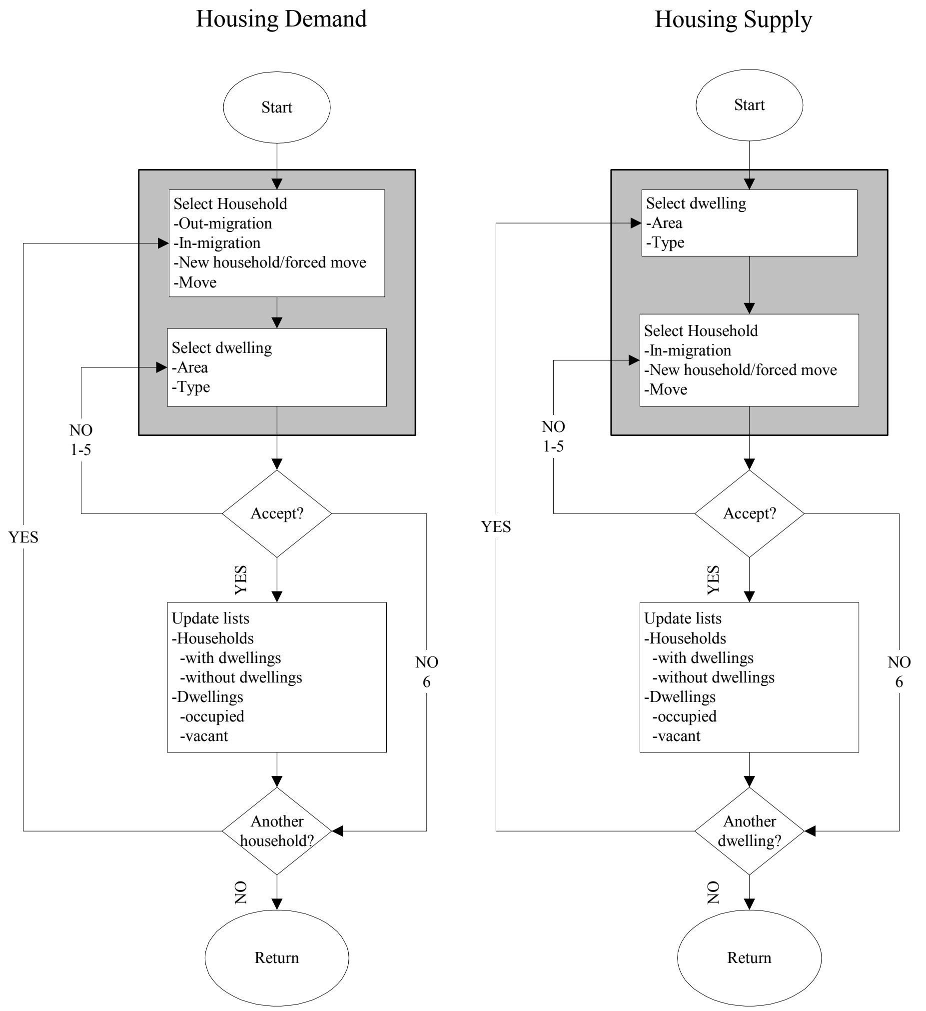


Figure 9: Microsimulatiion of household choice – ILUMASS

2.14 Summary of the models

This section provides an overview of the implemented models examined in this chapter. The models are compared with respect to the most important elements of residential mobility microsimulation models mentioned in section 2.5 aiming to point out their differences and similarities.

Elements to capture:	LOCSIM	HUDS	Fransson model	SVERIGE	UPDATE	HOUSIM
Database entities	Individuals Households Dwellings	Households Dwellings	Individuals Households Dwellings	Individuals Households	Individuals Households	Households
Space & Time dimensions	Only SPACE	Only SPACE	YES	YES	YES	YES
Occurrence of events	Monte Carlo sampling and decision rules	Probabilistic manner	Probability distributions, deterministic & stochastic rules	Monte Carlo sampling	List processing and Monte Carlo sampling	Monte Carlo sampling
Sequencing of events	Random order of events	Fixed order of events	Events occur for each individual only once a year	Fixed order of events	List processing	List processing
Event modules						
Demographic	Mortality Fertility Nest Leaving Divorce & Dehabitation Marriage of cohabitors Marriage & cohabitation Income change Non-triggered mobility In-migration Out-migration	Aging of the household head Changes in household size Changes in household income Changes in primary worker's job In-migration	Death Birth Nest Leaving Divorce Moving together Changing demands Income & education change In-migration Out-migration	Mortality Fertility Leaving home Divorce Cohabitation and Marriage Employment and Earnings Education Immigration Emigration	Mortality Maternity Divorce Marriage In-migration Out-migration	Aging of the household head Changes in household income

Elements to capture:	LOCSIM	HUDS	Fransson model	SVERIGE	UPDATE	HOUSIM
Housing Supply	Demolishment Splitting Combining Tenure change Renovation Rent & Market value update New housing construction	Demolishment Structure conversion Tenure change Renovation Rent & Market value update New housing construction	Housing expenditure All other changes in the housing supply are given exogenous	NO	NO	NO
Dimensions of housing search	Price Number of rooms Type Neighborhood quality Relative location Quality Age of house	Residential zone Minimal total accessibility cost for each workplace	Cost Size Attractiveness of the area Tenure Location	NO	Residential zone	Cost Tenure Mortgage amount Term of the mortgage
Housing Search Method	A probabilistic heuristic search model - Decision Plan Nets (DPNs) in conjunction with Monte Carlo techniques	Linear programming algorithm. Households are matched to dwellings in such a way that the total accessibility costs are minimized for each housing sub-market.	Rule of domination is used to structure decision making of the households	NO	A probabilistic matching of housing characteristics to housing needs.	New dwelling characteristics: 1) house price for already owner-occupiers and for first time buyers 2) mortgage amount 3) term of the mortgage
Calibration	YES	YES	YES	-	-	YES
Validation	YES	YES	YES	-	-	YES
Definition Year	1995-1997	1985	1994	1999	1995	1989

Table 1: Summary of implemented microsimulation models

Table 1 shows that although models do not take into consideration all elements and parameters of residential mobility, LOCSIM and Fransson models have the most complete definition and development. While both models seem to model all demographic changes, LOCSIM does not take into consideration the work & education related changes of individuals and the Fransson model leaves out several desirable dimensions of housing search. Additionally, the Franson model incorporates both the space and time dimensions whereas LOCSIM incorporates only the space dimension. On the other hand, LOCSIM is the most well documented model since the data flow diagram technique was used for the graphical representation of the model during design.

It should be mentioned that, although all models incorporate the spatial dimension of population, none of them provides spatial representation of model results at various levels of aggregation i.e. at residential zone and building block levels. The importance of the systematic handling of space in residential mobility model is evident. The lack of spatial representation is due to limitations imposed by the implementation platform chosen for each model.

Moreover, none of the models uses a formally defined technique for specifying, visualizing, and documenting the models during the phases of system development life cycle. The lack of using such a technique leads to models that are not comprehensible to all actors involved in the development of such a system and particularly in the system's end users i.e. the policy makers and consequently can not be easily extended with additional functionality by another development team.

Therefore, it is evident that there is the need to define a model that encompasses all elements of residential mobility. In addition it is important to use powerful computer science tools to develop and implement the model. These tools should allow designers and developers to develop a residential mobility system incrementally, that is, to be able to successively encompass new elements of residential mobility (new modules/sub-systems) without having to make major changes to the system already implemented. The *Unified Modeling Language (UML)* is such a tool, it will be used within this work and is presented in the following chapter.

2.15 Summary

In this chapter, the research conducted on residential mobility process was discussed and existing residential mobility microsimulation models were presented. Initially, the factors thought to influence residential mobility were discussed as well as the most significant results that have been accumulated over the last forty years. Then, the different strategies employed in the modeling of the residential mobility process were presented. Next, the microsimulation technique was presented and its strengths were pointed out. Then, the key concepts of a residential mobility microsimulation model were defined. Next, the LOCSIM, HUDS, Fransson model, SVERIGE, ILUMASS land-use model, UPDATE, HOUSIM and ILUTE residential mobility model were examined, and their main strengths and weaknesses were pointed out. Finally, an overview of the models presented was given and the contribution of this work in the field of residential mobility systems was outlined.

Chapter 3
Background Literature
in System Development Methods

3.1 Introduction

In this chapter, powerful computer science tools i.e. models, methods and processes are discussed which allow the effective conceptual and graphical representation of a system under development such as a residential mobility system. Section 3.2 presents the kinds of software system development methods where the predominance of object-oriented technology is pointed out. Section 3.3 presents the elements of object-oriented methods. Section 3.4 presents the Unified Modeling Language (UML), which is defined by the Object Management Group (OMG), as *the standard object-oriented* modeling language and the use of UML in microsimulation model development is illustrated, through an example: applying UML to the existing model LOCSIM. Section 3.5 defines the spatial, temporal and spatiotemporal concepts of systems. Sections 3.6 to 3.8 discuss two existing Spatio-Temporal modeling languages based on UML called Extended SpatioTemporal UML and SpatioTemporal PVL for UML as well as the ArcInfo Geodatabase Model, and their main strengths and weaknesses are pointed out. Moreover, in section 3.9 an overview of the modeling languages presented is given and the choices made within the context of this work are justified. Finally, section 3.10 defines ST-UML process, that is, a set of activities that guides the development of a system, by applying the UML modeling techniques extended with spatiotemporal constructs.

3.2 System development methods

The field of computer science has defined several methodologies for the development of software systems. Most of them may be categorized as being one of two kinds: *function/data* and *object-oriented*. Some of these methods have been used in the development of demographic and urban microsimulation systems.

Until recently, the development of such systems was based on a function/data analysis and design method, that is, the data flow diagram technique, in conjunction with a procedural programming language, such as FORTRAN and C. These methods enable the modeling of a system as a collection of functions that communicate through the exchange of data. However, there was a growing dissatisfaction with this approach due to the fact that these methods they do not seem to handle well the needs of complex systems, and are thus not well-suited for microsimulation applications [Miller97].

As a result, in recent years, a growing number of demographic and urban microsimulation systems are being developed within another framework called object-oriented. Object-oriented methods represent some of the most promising ways of meeting the demands of the most advanced applications [Miller97]. The main building block of an object-oriented system is the *object* (entity). Every object has identity (it can be named or otherwise distinguished from other objects), state (some data that define characteristics associated with it), and behavior (actions that affect the object's state and its relationships with other objects). Object-oriented analysis and design methods enable the modeling of a system as a collection of objects from the vocabulary of the problem domain.

The use of object-oriented technology offers several advantages over the function/data methods; however, seven of these can be considered the most important. These seven benefits include:

- *One-to-one mapping* of objects in the simulated world, to objects in the real world. This means that application objects can be manipulated in a way that is similar to the manipulation of real-world objects. This alone reduces the inherent complexity of the problem and increases the understandability of the system under development.
- *Modularity*: an object-oriented system may be decomposed into a set of individual components (modules) that are cohesive (strong connectivity among the elements of a single module) and loosely coupled (weak connectivity between the modules). The use of modularity is essential for managing the complexity of microsimulation applications. Several modules may be developed simultaneously and each module may be developed independently of the other ones.
- Objects, modules and even entire designs of a system under development may be *reused* in the development of other systems, leading to productivity improvements.
- The produced systems are more *resilient to change*, resulting in maintenance and cost benefits. A change will not affect the whole system; rather only one or more system modules will be affected (one or more objects).
- System development may be done in an *iterative* and *incremental* way. The system is developed and released gradually, and each release incorporates more functionality than the previous one, aiming to ultimately meet the requirements of the end user.
- The use of an object-oriented analysis and design method for the development of an information system, improves the collaboration between the different groups that participate in system development,

resulting in higher quality products. By matching real-world objects with system objects, analysts/ programmers can focus on understanding the problem rather than programming the solution. It is easier for end-users as well as for researchers/ analysts who do not possess computer programming skills, to understand the system under development, since object-oriented analysis and design is performed in the language of the problem domain and, therefore, does not contain implementation (code) details.

- Exploits the expressive power of object-oriented programming languages, resulting in productivity and code quality improvements.

3.3 The elements of object-oriented methods

Since the late 1980s object-oriented analysis and design methods have been highly popular and numerous methods have been developed (a comparison of 23 methods is given in (Monarchi and Puhr, 1992)). Most of them facilitate the development of a software system by providing the following mechanisms (Booch, 1994):

- *Notation*: the graphical representation of the method's concepts and semantics (model elements); the visual language used for the representation of the models of a system under development.
- *Process*: the activities that lead to the orderly construction of the system's model.
- *Tools*: the artifacts that facilitate model building and enforce rules on the models so that errors and inconsistencies of the model may be revealed.

The primary elements of an analysis and design method are the *concepts* underlying the method's notation, the *modeling techniques* for the graphical representation of the system under development, and the *process* the method provides (Hong *et al.*, 1993).

The framework of an object-oriented analysis and design method consists of the following concepts (Booch, 1994; Rumbaugh *et al.*, 1991; Svinterikou, 1998):

- *Object.* An object represents an entity that exists in time and space. An object has state, behavior and identity. The state and behavior of similar objects are defined in their common class. The *state* of an object encompasses all the properties of the object and their current values. The *behavior* is how an object acts and reacts, in terms of state changes and message passing. The *identity* of an object is the property that distinguishes it from all other objects. The terms object and instance are interchangeable.
- *Class.* A class is a set of objects that share a common structure and a common behavior.
- *Encapsulation.* Encapsulation is the process of compartmentalizing the state and behavior of a class. It serves to separate the interface of a class and its implementation. The *interface* of a class provides its outside view, that is, the operations and state variables of its objects. The *implementation* of a class comprises the representation of the operations defined in the interface of the class.
- *Association.* Association denotes a semantic relationship between classes, and consequently, a conceptual connection between instances of the classes participating in the association. The *multiplicity* of an association specifies how many objects of one class may relate to a single object of an associated class.
- *Aggregation.* Aggregation is a strong form of association. It is a "part-whole" relationship in which objects representing the parts of something

are associated with an object representing the whole (also called the aggregate).

- *Inheritance.* Inheritance is a relationship among classes wherein one class (*subclass*) shares the structure and/or behavior defined in one (*single inheritance*) or more (*multiple inheritance*) classes (*superclass*). Therefore, inheritance defines an 'IS A' relationship between classes where a subclass inherits all the behavior and properties of one or more superclasses. Typically, a subclass may define additional information and may refine the information it inherits to specialize it. An inheritance relationship is also called a *generalization relationship*.
- *Modularity.* Modularity is the decomposition of a system into a set of individual components (*modules*) that are cohesive (strong connectivity among the elements of a single module) and loosely coupled (weak connectivity between modules). Especially in large systems, the use of modularity is essential for managing complexity.

The modeling techniques a method provides have a profound influence upon how a problem may be attacked and how a corresponding solution may be shaped (Booch *et al.*, 1999). According to (Booch, 1994), it is impossible to capture all the details of a complex system in just one view of a model. That is why most methods provide more than one modeling technique for the graphical representation of the system under development. These techniques allow the modeling of a system from different viewpoints, each capturing important aspects of the system.

The process of a method guides the modeler during the construction of the system model. It consists of the following properties:

- *A set of steps* recommended by the method that cover (usually part of) the system development lifecycle. In particular, the process of a method defines the sequence of these steps, the required inputs and outputs, and the interaction between these steps.

- *Methodological guidelines and rules* that assist the modeler in following the process steps in the most effective way. They may refer to the identification of the various elements of the system under development, to the modeling of these elements, to the transformation of a model from a higher level of abstraction to a lower one (e.g. implementing a design model using a particular implementation platform).

According to (The Object Agency Inc., 1995), in addition to these primary process properties, the following properties are considered to be desirable features of a process:

- *Domain analysis* is the examination of existing related systems in order to identify the classes and objects common to all applications within a given domain (Booch, 1994). A method should consider domain analysis as part of its process activities.
- *Traceability* is the ability to trace objects in one model to objects in another model (Jacobson *et al.,* 1992). A process should provide guidelines to ensure traceability through the development life cycle.
- *Reuse* of software development artifacts such as classes, designs and code. A process should provide guidelines that facilitate artifact reuse.

Most methods provide processes that may be used in an *iterative* and *incremental* approach of development, that is, the software is developed and released gradually, and each release incorporates more functionality than the previous one, aiming to ultimately meet the requirements of the end user.

Until recently, the methods scene was split and competitive (Fowler and Scott, 1997): each method had its own group of practitioners, and although all of these methods were quite similar, basic concepts would appear in very different notations, causing confusion to the users of the methods. The need of a standard object-oriented modeling language emerged. In 1997 the Object Management Group (OMG) adopted the Unified Modeling Language (UML) as the standard object-oriented modeling language which is discussed in the next section..

3.4 Unified Modeling Language (UML)

The *Unified Modeling Language (UML)* is a language for specifying, constructing, visualizing, and documenting the artifacts of a software-intensive system. The main characteristics of UML are:

- It represents a collection of "best engineering practices" that have proven successful in the modeling of large and complex systems. UML fuses the concepts of the Booch, OMT and OOSE methods.
- It consists of a set of semantics and notation that addresses a wide variety of domains, life cycle stages and implementation technologies. Additionally, it provides a set of extension mechanisms that allows the core concepts of UML to be tailored according to the needs of specific applications and domains.
- Moreover, UML focuses on a standard modeling language, not a standard process. It offers a set of concepts and a set of graphical notations for representing these concepts, which is process independent. As a result, UML may be used in conjunction with the processes defined in Booch, OMT, OOSE and many other methods.
- Finally, UML is formally defined through a metamodel expressed in UML.

UML defines a set of modeling techniques that provide multiple perspectives of the system under development.

1. *Use case diagrams* show the behavior of the system under development without specifying its internal structure. It uses *actors* that are the roles played by users with respect to the system, and *use cases* where each one represents part of the system behavior that is performed during the interaction of an actor with the system.
2. *Class diagrams* show the static structure of the system, that is, the existence of *classes*, their *attributes*, *operations*, *relationships* with other classes, and the *constraints* that apply to classes and their properties. Classes and the relationships among them may be grouped into *packages.*
3. *Object diagrams* show the static structure of the systems just as do class diagrams, but from the perspective of real or prototypical cases. It uses *objects* and their *relationships* to show the static snapshots of instances of the information captured in class diagrams.
4. *Sequence diagrams* show groups of *objects* (example objects) and the *messages* exchanged among these objects. Typically, an interaction diagram traces the execution of a single use case. It emphasizes the sequence of the message exchange that takes place within a use case.
5. *Collaboration diagrams* capture the behavior of use cases in the same context as sequence diagrams. A collaboration diagram represents a *Collaboration*, which is a set of *objects* related in a particular context and an *Interaction*, which is a set of *messages* exchanged among the objects within the collaboration. In contrast to sequence diagrams, collaboration diagrams emphasize static connections between objects.
6. *Statechart diagrams* show all the possible *states* in which a particular object may be, the *events* that cause the *transition* from one state to another, and the actions that result from a state change.

7. *Activity diagrams* are a special case of statechart diagrams, in which the states (*action states)* are activities representing the execution of operations, and *transitions* are triggered by the completion of the operation executions. Additionally, an activity diagram uses *decisions* to indicate different potential transitions of an action state depending on some constraints.
8. *Component diagrams* show the structure of the code itself, that is, the dependencies among software components, including source code and executable components. A component diagram uses *components* to represent physical modules of code.
9. *Deployment diagrams* show the physical relationships among software and hardware components in the delivered system. A deployment diagram consists of *components, nodes* that represent pieces of hardware, and *connections* and *dependencies* to show the communication paths among nodes and components respectively.

The graphical notation of the elements used in UML diagrams is presented in **Appendix A**.

In addition, UML defines *three built-in extension mechanisms* that are common to all diagrams previously described, and apply consistently throughout the language. These mechanisms include tagged values, constraints and stereotypes, and are defined to meet the needs of applications that require additional features and/or notations beyond those defined in UML.

1. A *tagged value* is an extension of the properties of a UML element, allowing the creation of new information in that element's specification.
2. A *constraint* is an extension of the semantics of a UML element, allowing the addition of new rules or the modification of existing ones.
3. A *stereotype* is an extension of the vocabulary of the UML, allowing the creation of new kinds of building block similar to existing ones but

specific to the problem domain. In other words, a stereotype may introduce additional properties (tagged values), additional constraints and a new graphical representation into the UML element to which it is attached. An example of a stereotype predefined in UML is the <<utility>> stereotype, which may be attached to the class element. A *utility stereotyped class* is a kind of class, which is used to group the attributes and/or operations that may be accessed by more than one class. Utility classes do not have objects; the class itself serves as the sole instance. Figure 10 shows the graphical representation of the MathFunction utility class.

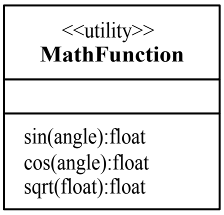


Figure 10: Utility class

In order to illustrate the use of UML in microsimulation model development, UML is applied at the *conceptual phase* of an existing model called LOCSIM (Oskamp, 1995; Oskamp, 1997). LOCSIM as discussed in **Chapter 2** is a microsimulation model that captures the developments in demographic structures and housing markets, as well as the interrelation between them. It was developed using conventional analysis and design methods, and operates at local levels and for short periods of time within a time horizon of three to five years.

The model consists of three main sub-models: *demographic*, *housing supply*, and *housing market*. Within the model, the *individual*, *household* and *dwelling* entities (classes) are of interest. Figure 11 shows these classes, their attributes and the relationships among them, using the *class diagram* UML modeling technique. In LOCSIM:

- *Individual* class keeps information concerning the sex, education, job, income and marital status of each individual. An individual may be either a household head (HEAD=1) or a household member (HEAD=0). A household head is associated with zero or more household members.
- *Household* class keeps information concerning the size and total income of the each household. A household has one household head and zero or more household members. A household lives in a dwelling.
- *Dwelling* class keeps information concerning each dwelling such as its type, number of rooms, location and ownership. A dwelling may be occupied by one household or be unoccupied.

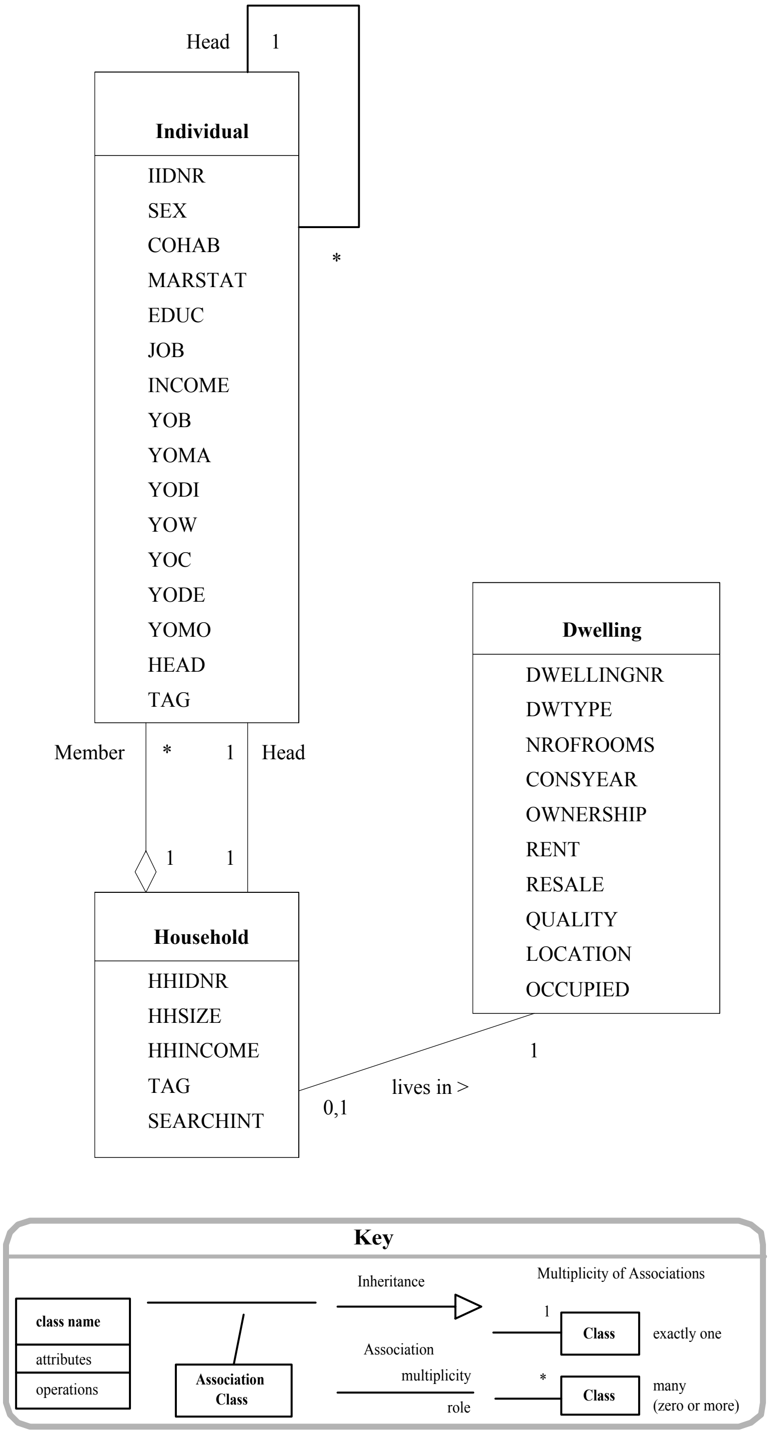


Figure 11: LOCSIM class diagram

Monte Carlo sampling is used to determine the occurrence of different events, which include: *mortality*, *fertility*, *nest leaving*, *divorce/dehabitation*, *marriage/cohabitation*, and *mobility*. Individuals are sequentially exposed to the risk of experiencing these events. The occurrence of different events may change the demographic structure of a household, which may in turn increase a household's desire to move. LOCSIM keeps track of the household desire to move, which may lead to different types of mobility: residential migration, out-migration or in-migration. Within LOCSIM, mobility is either the direct result of *household formation and dissolution*, or is based on *age-specific mobility rates*. As a household experiences a demographic change, it is assigned a move "tag" to indicate that the given household is searching for a new residence (*TAG* attribute of *Household* class). A tag can be a *preferred* move tag or an *implied* move tag. All households do not have an equally strong desire to find a new residence. For example, a household wanting to move to a new dwelling because of the birth of a child, will have less of a desire to move compared to a divorcing couple. This difference is captured by introducing *search intensity* (*SEARCHINT* of *Household* class). Figure 12 shows the *use case* view of LOCSIM, that is, the model's desired behavior. In LOCSIM demographic events are modeled as distinct use cases. These use cases incorporate some common behavior, which is the effectuation of the desired move, modeled as a distinct use case called *Moving Behavior*.

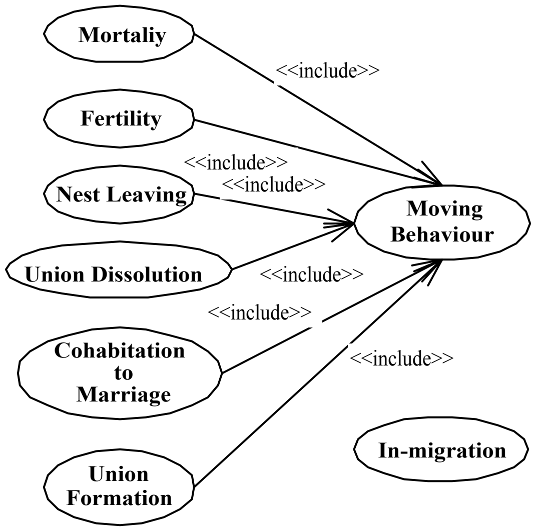


Figure 12: LOCSIM use case diagram

LOCSIM use cases are further elaborated through scenarios. In addition, each use case scenario is further elaborated through sequence diagrams. Each sequence diagram shows the objects that participate in a use case scenario and their behavior within the use case. In this example, two LOCSIM use case are further elaborated: *Fertility* and *Nest Leaving*.

Fertility use case scenario

- All women aged 18 to 45 are subjected to a probability of giving birth.
- If a child is born, a new individual is created and added to the population list. Using Monte Carlo sampling, the sex of the child is determined.
- The birth leads to an increase in household size.
- The household is tagged with a preferred move tag.

Figure 13 shows the *Fertility* sequence diagram.

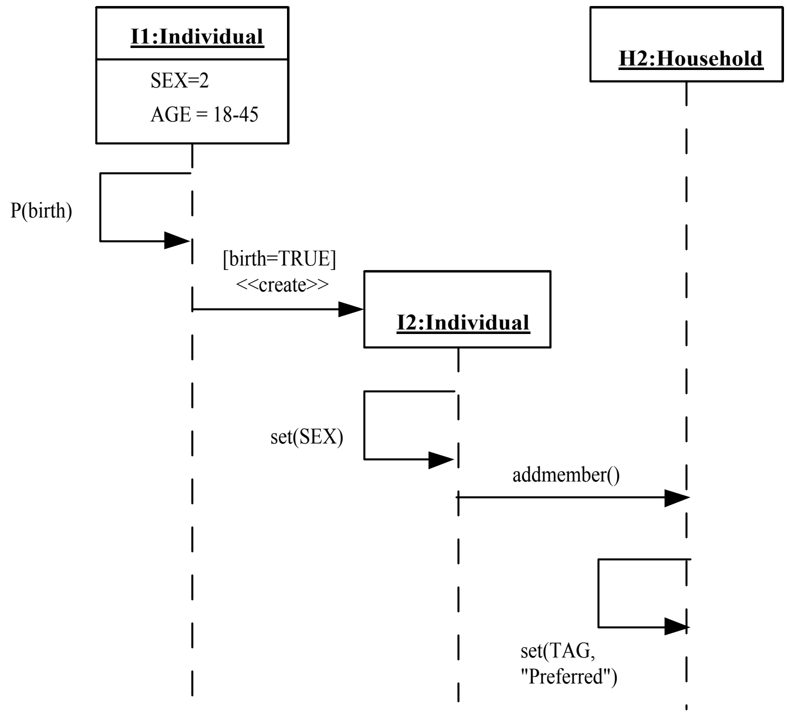


Figure 13: Fertility sequence diagram

Nest Leaving use case scenario

- All individuals of age 18 and above, who are living with their parents, are subjected to a nest leaving probability.
- The original household decreases in size and is therefore tagged with a preferred move tag.
- The individual simulated to leave its parents, forms a new single-member household, tagged with an implied move.
- A dwelling search takes place, which is modeled in *Moving Behavior* use case.
- If the individual's house search is unsuccessful, the individual returns to the original household and the new household no longer exists. The original household no longer has a preferred move tag.
- If the individual's house search is successful the move actually takes place, and the new household no longer has an implied move tag.

Figure 14 and Figure 15 show the *Nest Leaving* sequence diagrams.

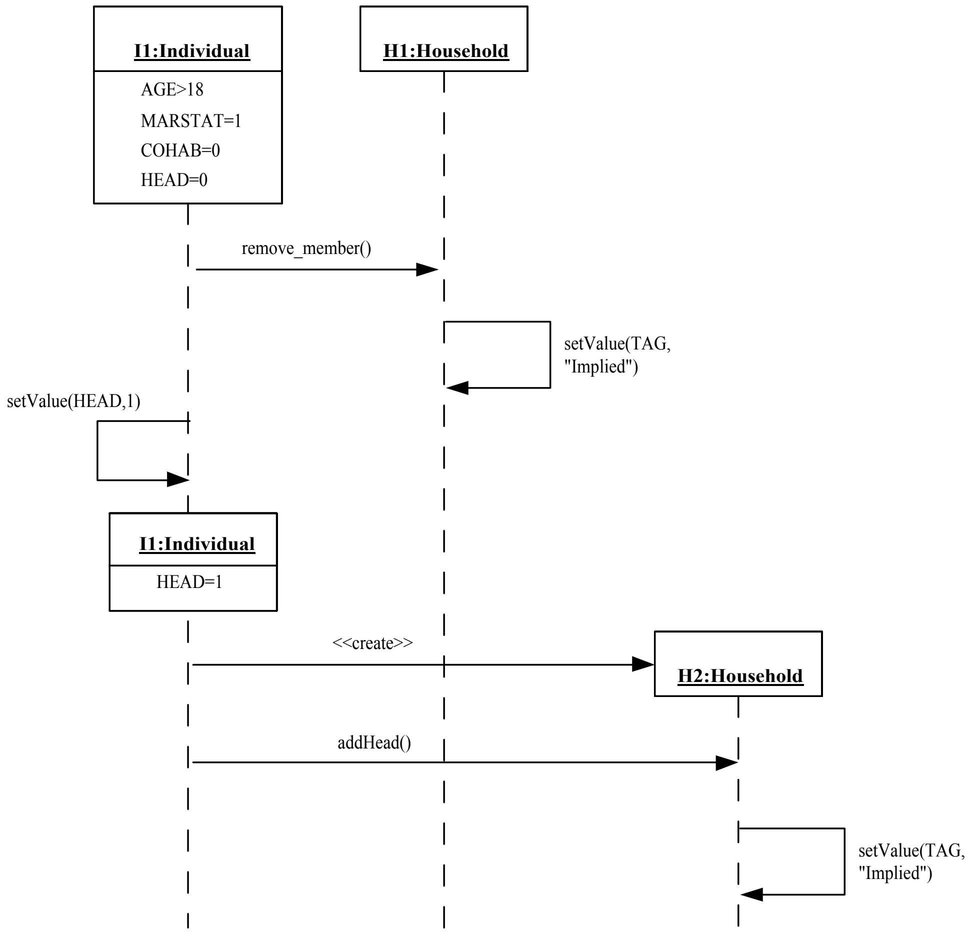


Figure 14: Nest Leaving sequence diagram I

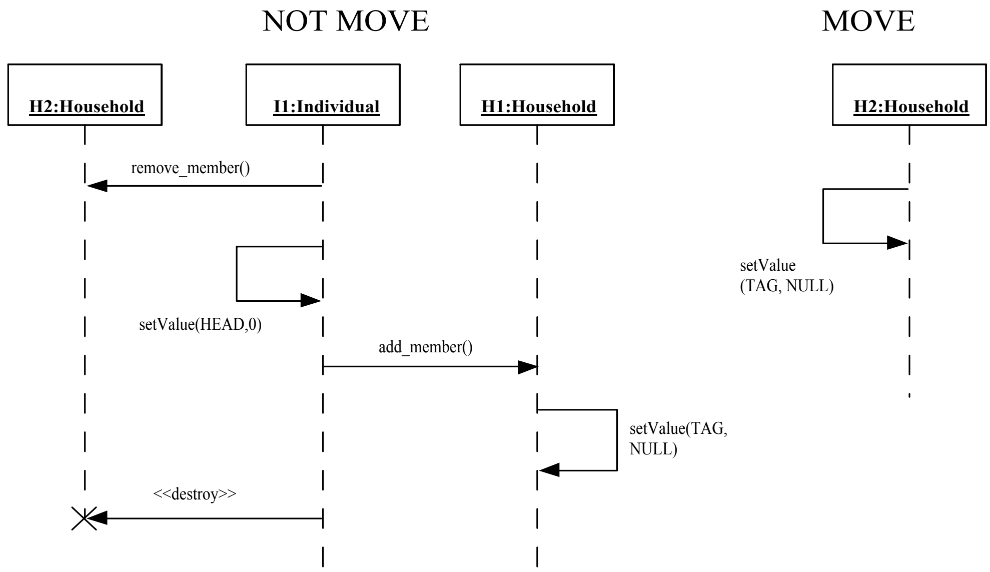


Figure 15: Nest Leaving sequence diagram II

Further elaboration of LOCSIM class diagram and sequence diagrams lead to diagrams whose mapping to code in an object-oriented language is quite straightforward either for a programmer, or for a forward engineering tool.

3.5 The concepts of spatiotemporal methods

The ability to capture and analyze population changes over time and to incorporate the spatial dimension of population is of great importance to a residential mobility model, as outlined in **Chapter 2**. That is why, the methods used to develop a residential mobility system should be able to capture spatial, temporal and spatiotemporal concepts.

3.5.1 Spatial concepts

The key concepts of a system incorporating spatial information are (Pfoser and Tryfona, 1998; Tryfona *et al.,* 1997; Tryfona and Jensen, 1999; Price *et al.*, 1999b):

- *Space.* A set of elements called points; finite sets of *points* (i.e. subsets of space, which can be point, lines or regions) are called *geometric figures*. Space may be defined as a set of any kind of elements, even non-numeric ones. However, the Euclidean space, which is based on a coordinate system allowing standardized metric measurements, is used in current spatial applications.
- *Space Views.* There are two competing views of Euclidean space (Price *et al.*, 1999b):
  - *Object View:* discrete, identifiable objects associated with spatial and thematic properties.
  - *Field View*: a framework of spatial locations across a field where thematic data is measured (e.g. temperature or different values of soil type across an area).
- *Model of Space*.
  - The *continuous model* that views space as isomorphic to real numbers.
  - The *discrete model* that views space as isomorphic to the integer numbers.
- *Space Dimensions.* Spatial dimensionality is the number of axes or coordinates considered. Both the dimensions of the underlying space and the information being modeled within that space should be considered, e.g. 1D points and 2D polygons in 3D space.
- *Spatial Information.*
  - *Spatial Objects.* Real-world objects have a *position* in space. An object whose position in space (*spatial extent*) is of interest within an information system, is called a spatial object (e.g. the position of a *dwelling* is of interest within a residential mobility system). *Position* is fully and non-redundantly defined by four components: *shape*, *size*, *location (centroid)* and *orientation* (Tryfona and Jensen, 1999).

- *Spatial Attributes.* Spatial objects may have attributes whose value is a spatial extent (e.g. the *soil type* of a land parcel or the *walk distance zone* of a dwelling).
- *Spatial Relationships*. Spatial objects may be related to each other in space via spatial relationships. Spatial relationships among objects are in fact relations on the objects' positions. There are three kinds of spatial relationships (Tryfona and Jensen, 1999): *topological* (e.g. INSIDE, OUTSIDE), *directional* (e.g. NORTH OF) and *metric* (e.g. 5KM AWAY FROM).

3.5.2 Temporal concepts

The key issues involved in the analysis and design of a system incorporating temporal information are (Theodoulidis and Loucopoulos, 1991):

- *Model of time*: In temporal databases three basic time models have been proposed (Snodgrass, R., (ed), 1995):
  - The *continuous model* that views time as isomorphic to real numbers.
  - The *dense model* that views time as isomorphic to the rational numbers.
  - The *discrete model* that views time as isomorphic to the integer numbers.

  In all three models, a point on the underlying time axis is called a *time point*[2]; the time between two time points, that is, a duration of time with a starting and ending point, is called a *time period*[3]; an unanchored duration of time is called an *interval*.
- *Time linearity*: Time axis may be considered as *linear, periodic* or *branching.* In the linear model time advances from the past to the present and then to the future in a totally ordered form. The periodic

[2] In the consensus glossary it is called *instant* [Jen94].

[3] In the consensus glossary it is called *interval* [Jen94].

model allows the modeling of requirements, which refer to the periodicity of time units such as weeks, months etc., and of user-defined periods such as weekends, work-weeks etc (Theodoulidis and Loucopoulos, 1991). In the branching model time is linear from the past up to the present, where 3it branches into several time lines each one representing a potential sequence of events.

- *Time boundaries*: Conceptually, time may extend from the infinite past to the infinite future. However, in many temporal systems time is bounded at both ends, as unbounded time requires variable sized data structures for its storage (Snodgrass, R., (ed), 1995).
- *Time reference*: Time is called *absolute* if it does not depend either on the current time or on other times, and it is called *relative* if it references some other time. An example of absolute time is 'October 10, 1972' and an example of relative time is 'last Christmas'.
- *Calendars and granularities.* A *calendar* provides a human interpretation of time (Jensen *et al.,* 1994). As such, calendars determine the mapping between human-meaningful time values and an underlying time-line. The calendar supported by most temporal systems is the *Gregorian calendar*, defined in (Jensen *et al.,* 1994). Calendars define granularities: *granularity* is a partition of the time axis into a finite number of discrete segments. The default granularities of the Gregorian calendar are the *year, month, day, hour, minute* and *second.* Different temporal applications require different levels of granularity. Therefore, a temporal model should enable the user to choose the appropriate granularity for his/her application.
- *Object identification across time*: In systems that incorporate temporal information, object identification takes on an added dimension as it is necessary to recognize that an object currently under examination is in fact the same as some other object already known to the database

(Theodoulidis and Loucopoulos, 1991). Thus, a *time-invariant identifier* is needed, whose purpose is to ensure that a single real-world object is represented only once in the system, together with all of its time-varying properties.

- *Time dimensions*: Three time dimensions are of general interest: user-defined time, valid time and transaction time (Jensen *et al.,* 1994).
    - *User-defined* time is an uninterpreted attribute domain of date and time. It is parallel to domains such as "money" and integer, and it may be used for attributes such as "birth day" and "hiring date".
    - The *valid time* of a fact is the time when the fact is true in the modelled reality. A fact may have associated any number of *time points* and *time periods*, with single time points and periods being important special cases. Valid times are usually supplied by the user.
    - The *transaction time* of a fact is when the fact is current in the database and may be retrieved. Transaction times are consistent with the serialization order of the transactions. Transaction-time values cannot be later than the current transaction time. In addition, as it is impossible to change the past, transaction times cannot be changed. Transaction times are system-generated and supplied.

  A system that supports both valid and transaction time is termed *bitemporal*.
- *Timestamping*: A *timestamp* is a time value associated with some information. The timestamp may be of valid time, transaction time or both. A timestamp may be represented by a time point, a period etc. expressed in a particular granularity (*timestamp granularity*). The granularity of the timestamp denotes the precision to which the timestamp can be represented (Kakoudakis, 1996). Two approaches have been proposed for timestamping: *property timestamping* and *object timestamping.* In property timestamping, a timestamp may be associated

with each time-varying attribute of an object, while in object timestamping, a timestamp may be only associated with the whole object.

3.5.3 Spatiotemporal concepts

The spatial and temporal concepts described in the previous sections are combined to produce the following spatiotemporal concepts (Price *et al.*, 1999b, Price *et al.*, 2000):

- *Objects (Classes)*
    - *Spatial object (class)*: the object's position in space, i.e. spatial extent, is of interest.
    - *Temporal object (class)*: the object associated with one or more timestamps, representing the object’s valid and/or transaction time.
    - *Spatiotemporal object (class)*: an object associated with a spatial extent and one or more timestamps representing the object’s valid and/or transaction time.
    - *Temporally dependent spatial object (class)*: an object with a spatial extent associated with one or more timestamps, representing the spatial extent’s valid and/or transaction time.
    - *Temporally dependent spatiotemporal object (class)*: an object associated with one or more timestamps, representing the object’s valid and/or transaction time and a spatial extent associated with one or more timestamps, representing the spatial extent’s valid and/or transaction time.
- *Attributes*
    - *Spatial attribute*: the attribute value is a spatial extent.
    - *Spatially dependent thematic attribute*: the attribute has a set of thematic values, each associated with a spatial extent representing

the location where that attribute value is valid, i.e. the attribute values may change over space and their changed values are retained.

- *Temporally dependent thematic attribute*: the attribute has a set of thematic values, each associated with one or more timestamps (the value's valid and/or transaction time), i.e. the attribute values may change over time and their changed values are retained.
- *Temporally dependent spatial attribute*: a spatial attribute whose value is associated with one or more timestamps representing the spatial extent's valid and/or transaction time.
- *Spatiotemporally dependent thematic attribute*: a spatially dependent thematic attribute that has a set of thematic values, each associated with a spatial extent and one or more timestamps.

- *Associations*
  - *Spatially dependent association*: each link of the association is associated with a spatial extent representing the location where the link is valid, i.e. the link may change over space and its changes are retained.
  - *Temporally dependent association*: the association has a set of links, each associated with one or more timestamps (the value's valid and/or transaction time), i.e. the links may change over time and their changes are retained.
  - *Spatiotemporally dependent association*: the association has a set of links, each associated with a spatial extent and one or more timestamps.

The spatial, temporal and spatiotemporal concepts presented in this and the preceding sub-sections are not incorporated in the set of the UML core elements (Price *et al.*, 2000). However, UML provides a set of *extension mechanisms* to allow core concepts to be tailored according to the needs of spatiotemporal applications. Two object-oriented modeling languages have been proposed as extensions of UML with spatial, temporal and spatiotemporal semantics and notation. In addition, an object-oriented modeling language has been proposed by ESRI as an extension of UML with spatial semantics and notation. These modeling languages are presented in the following sections.

3.6 Extended SpatioTemporal UML

The *Extended SpatioTemporal UML (ESTUML)* (Price *et al.*, 2000; Price *et al.,* 1999a; Price *et al.*, 1999b; Price *et al.*, 1999c), is *a general conceptual data modeling language* suitable for specification and analysis of spatiotemporal database applications. The main characteristics of ESTUML are:

- It is an extension of UML, a modeling language with a high level of acceptance, understandability and the flexibility needed for extensions.
- It is a generic language that can support a range of spatiotemporal data models and types, and can be mapped to different implementation systems.
- ESTUML's semantic modeling constructs and their corresponding notation are formally defined through functional specification.
- The latest version of ESTUML does not use the UML extension mechanisms; rather, *the constructs defined to extend UML with spatiotemporal semantics, go beyond the extension mechanisms available in UML.* It will be difficult to incorporate these extensions, since most UML tools allow the extension of their modeling constructs only though UML extension mechanisms.

Within Extended SpatioTemporal UML, UML is extended with the addition of five new symbols named *Spatial*, *Temporal*, *Thematic*, *Group* and

Existence-Dependent, and a *Specification Box* describing the detailed semantics of the spatiotemporal data represented by the new symbol.

- The *spatial* symbol represents a spatial extend, which consists of an arbitrary set of points, lines, regions or volumes. The spatial extent may be associated with thematic or composite data, or may be used to define an attribute domain.
- The *temporal* symbol represents a temporal extent, or timestamp, which may be associated with thematic, spatial or composite data. Timestamps may represent existence time for objects, valid time for associations and attributes, and transaction time for objects, associations, and attributes.
- The *thematic* symbol represents thematic data. It can only be used at the attribute level and only in conjunction with one of the other two symbols to describe an attribute with temporal or spatial properties. A thematic attribute domain with no spatial or temporal properties uses standard UML notation.
- The *group* symbol is used to group attributes with common spatiotemporal properties or inter-attribute constraints. It is used to graphically illustrate associated sets of attributes and avoids the possibility of redundantly specifying the same spatial extents or timestamps for attributes.
- The *existence-dependent* symbol is used to describe attributes and associations dependent on object existence, i.e. they are defined only when the related object(s) exist.
- Each object, association, attribute or attribute group may be associated with a *Specification Box* to provide details of its spatiotemporal semantics. A specification box includes information on the time units (e.g. instant, interval), dimensions (e.g. valid, transaction), model (e.g. linear, periodic) and interpolation (e.g. discrete, step), and space dimensions (0D, 1D or 2D), model (object or field based) and

interpolation. Specification Boxes of a single class (or association class) are collected in an extra specification compartment and added to the class.

The first three constructs (spatial, temporal and thematic) can be applied at different levels of the UML class diagram and in different combinations to add spatiotemporal semantics to a UML model element.

- At the attribute (and attribute group) level, spatially, temporally, and spatiotemporally dependent thematic attributes, spatial attributes, and temporally dependent spatial attributes may be modeled.
- At the object class level, spatial, temporal and spatiotemporal objects (classes), and temporally dependent spatial and spatiotemporal objects (classes) may be modeled.
- At the association level, spatially, temporally and spatiotemporally dependent associations may be modeled.

A simple example is used to illustrate the STUML notation. It concerns information about household employment and is shown in Figure 16**.** In this example:

- A *household* class has an income, a family size, a household head age and it is denoted whether it is a new household or not. Its objects' existence time is of interest.
- Changes of the *income* and *family size* attributes' values over time should be stored in the database.
- A *job position* class has a position (spatial extent) that may change over time. Its objects' existence time is of interest. The class has the IndustryType, Occupation and WorkplaceZone attributes.
- *WorkplaceZone* is a spatial attribute, whose values may change over time (spatiotemporal attribute).

- The *Primary Worker has* association between *Household* and *Job Position* classes may change over time.

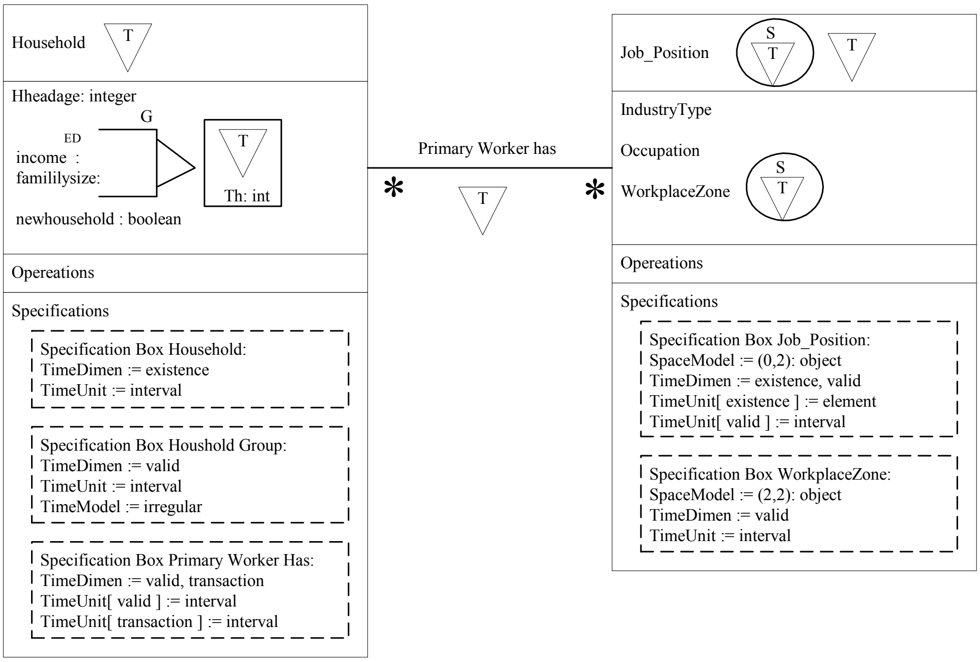


Figure 16: An example in Extended SpatioTemporal UML (ESTUML)

The major weakness of the modelling language is that there has not been developed yet a visual modelling tool that supports ESTUML. Since ESTUML does not use the UML extension mechanisms to define spatiotemporal concepts, it cannot be defined through the UML metamodel and the Object Constraint Language (OCL). Consequently, tool developers will have difficulty in incorporating these spatiotemporal extensions in their UML tools.

3.7 SpatioTemporal PVL for UML

The *SpatioTemporal PVL for UML* (ST-PVL for UML) (Bedard, 1999; Brodeur *et al.*, 2000; Bedard and Proulx, 2000) is a visual modeling language suitable for spatiotemporal database modeling. It involves two aspects:

1. The Spatiotemporal Plug-in for Visual Languages (ST-PVL). It is a generic set of spatiotemporal concepts that allows the effective modeling of spatiotemporal information and may be applied to create a spatiotemporal extension for *any* visual modeling language. The main characteristics of ST-PVL are:
   a. It allows for independence of the solution from any visual modeling language while remaining pluggable into the users' favorite language and tool.
   b. It targets the level of abstraction of conceptual database modeling (i.e. an analysis level closer to the client's view than usual) to facilitate analysis/conceptual spatial and temporal modeling without bothering with implementation issues.
2. The Unified Modeling Language (UML). It is the visual modeling language into which the ST-PVL is integrated. Within UML, the ST-PVL concepts are defined as a set of spatial and temporal stereotypes that apply to the Class diagram.

A visual modeling tool, named *Perceptory*, is developed to support ST-PVL for UML. It allows users to create conceptual spatial and spatiotemporal object class models, document these objects using an object class and project dictionary, automatically produce various reports on the content of the dictionary, and generate the skeleton programming code for different implementation platforms (e.g. Oracle8i, MapInfo). In addition, Perceptory is based on a repository especially developed for spatial and spatiotemporal databases. Finally, it is developed on a stencil supported by the Visio software (v. 5.0 or 2000) and conforms to the international spatial database standards ISO TC-211 and OGC.

ST-PVL for UML is formally defined through the Perceptory's geospatial data repository metamodel. The metamodel uses UML object class metamodel elements, and integrates ISO TC-211 and OGC necessary elements.

ST-PVL for UML uses five basic constructs named *point* (0-dimensional geometry), *line* (1-dimensional geometry), *area* (2-dimensional geometry), *punctual time* (0-dimensional temporality), *and durable time* (1-dimensional temporality). The fist three constructs can be combined in four different ways (simple, alternative, complex, multiple) with cardinalities in order to properly represent spatial properties of objects (classes), attributes and associations.

- *Simple geometry*. The geometry of an object may be a point, line, or area (non-decomposable geometry).
- *Complex geometry*. It is a spatial definition with more than one spatial components that, when taken together, constitute the geometry of an object (e.g. a park delineated by an aggregate of areas). It can be constructed from 0D, 1D or 2D primitives exclusively, or combinations of them.
- *Multiple geometry*. It is a spatial definition comprised of more than one simple, complex and/or alternate spatial components. All components of multiple geometry shall be instantiated and all have specific meanings (e.g. a *city* object may be represented by an area and a point denoting its downtown location).
- *Alternate geometry.* It is a spatial definition comprised of more than one simple, complex and/or multiple spatial components. However, only one geometry shall be instantiated at a time (e.g. a *river* will be shown as a line if its width is less that 10 meters, otherwise it will be shown as an area).

Punctual time and durable constructs can be combined in four different ways (simple, alternative, complex, multiple) with cardinalities in order to properly represent temporal properties of objects (classes), attributes and associations. These combinations of temporality constructs have the same semantics with the ones defined for geometric constructs. Moreover, they may be applied to spatial attributes, spatial associations and the geometry of objects (classes) to define the *spatial evolution* of attributes, associations and classes respectively over time. Details of the spatial, temporal and spatiotemporal semantics are defined in a *formal dictionary.*

The household employment example presented in the preceding section is used to illustrate the ST-PVL for UML notation and is shown in Figure 17**.**

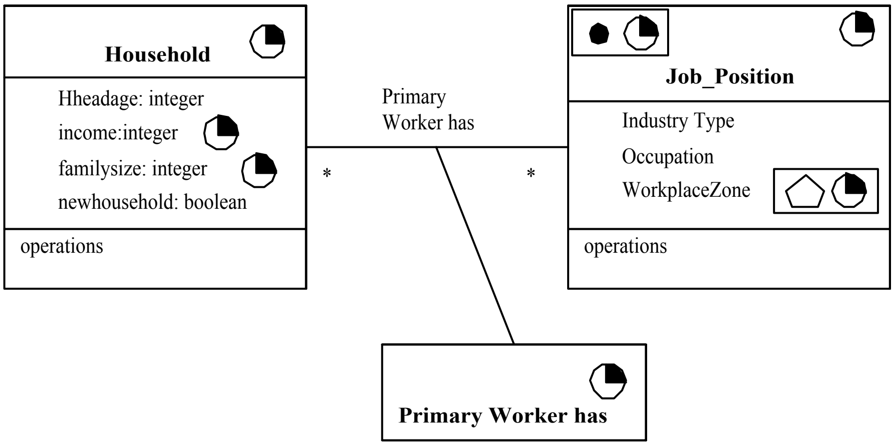


Figure 17: An Example in SpatioTemporal PVL for UML

3.8 The ArcInfo UML Geodatabase Model

ESRI ArcGIS defines a new object-oriented data model called *ArcInfo geodatabase model* which brings the modelling and implementation of geographic features closer to their actual real-world counterparts (ESRI, 2001). Therefore, one may create a *geodatabase* (geographic database) by defining parcels, dwellings, neighbourhoods as well as a number of characteristics for each, such as fields, validation rules, relationships and subtypes.

The geodatabase data model supports, as standard, a rich collection of objects (rows in a database table) and features (objects with geometry). The model supports advanced *feature types* such as geometric and logical networks, true curves, complex polylines, and user-defined features. Vector features can have two, three, or four *dimensions* (x, y, z and m).

In addition, one can define topological and associative relationships and rules that define how feature classes interact. A relationship is an association between two objects. These objects can be non spatial (objects) or spatial (feature). Besides identifying the associated objects relationships can have additional properties.

Moreover, the model supports the definition of *sybtypes* and *domains* to ensure the attribute integrity of the database:

- *Subtypes* allow the developer to define categories for features, which allow him/her to distinguish objects base on different behaviours, attributes and access privileges. For example, if telephone poles only have three locations (highways, arterial roads, secondary roads), the user subtypes can be set for each pole location so that during data entry, users can easily associate the proper attribute.
- *Domains* control the allowable set of values that an attributes may have. There are two types of attributes domains:
    - *Coded Values* specify listings of acceptable values. For example, the type of telephone poles could be coded as 1 (Wood), 2 (Steel) or 3 (Cable).

- *Range* specifies acceptable starting and ending numeric values. For example, telephone pole heights could be between 20 and 50 feet.

The Computer-Aided Software Engineering (CASE) tools such as Microsoft Visio, allow developers to create design models of the geodatabase structures using UML. The ArcInfo UML Model diagram is a set of Visio Drawing templates defined by ESRI and contains the object model required for modeling a geodatabase using UML in Microsoft Visio. It consists of five packages (Figure 18): 1) *Logical View*, 2) *ESRI Classes*, 3) *Workspace*, 4) *ESRI Interfaces* and 5) *ESRI Network*, which act as directories where different parts of the entire model are maintained. The Logical Package is the root level and contains the other packages. Database designers and developers can use the Workspace package to create their designs. The ESRI Classes package contains classes (e.g. *Object* and *Feature*) that extend UML concepts to represent components that incorporate a spatial dimension and access spatial data sources, including geodatabases. Classes defined by the designer will inherit from these classes. The remaining packages are not used in database design. In addition, the ArcInfo UML Model diagram defines a set of constraints and validation rules that should apply to UML constructs (e.g. many –to-many relationships are not allowed).

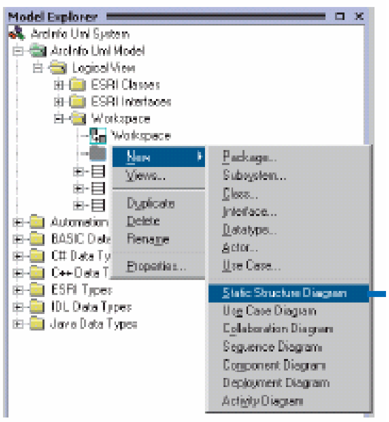


Figure 18: ArcInfo UML Model diagram packages

Once the model is created, it may be exported to the Microsoft Repository by Microsoft Visio. Then, the Schema Wizard in ArcCatalog (ArcGIS) may be used to create the database schema.

3.9 Summary of the modeling languages

This section provides an overview of the modeling languages examined in this chapter. The modeling languages are compared with respect to the most important elements of spatio-temporal methods mentioned in section 3.5 aiming to point out their differences and similarities.

Elements to capture:	Extended SpatioTemporal UML	SpatioTemporal PVL for UML	ArcInfo Geodatabase Model
Space	YES	YES	YES
Space Views	Object - Field	Object - Field	Object
Model of Space	Continuous – discrete		
Space Dimensions	1D,2D,3D	1D,2D,3D and various combinations of them	1D,2D,3D
Spatial Information	Spatial Objects Spatial Attributes Spatial Relationships	Spatial Objects Spatial Attributes Spatial Relationships	Spatial Objects
Model of time	Continuous, Dense, Discrete		NO
Time linearity	Linear, periodic, branching		
Time boundaries	ANY	ANY	
Time reference	Absolute, relative		
Calendars and granularities	Gregorian	Gregorian	
Object identification across time	YES	YES	

Elements to capture:	Extended SpatioTemporal UML	SpatioTemporal PVL for UML	ArcInfo Geodatabase Model
Time dimensions	User-defined Valid time Transaction time	User-defined Valid time	
Timestamping	Property timestamping	Property timestamping	
Spatiotemporal Information			
Attributes	Spatially, temporally, and spatiotemporally dependent thematic attributes Spatial attributes Temporally dependent spatial attributes	Spatial, temporal and spatiotemporal attributes	
Classes (Objects)	Spatial, temporal and spatiotemporal objects (classes) Temporally dependent spatial and spatiotemporal objects (classes)	Spatial, temporal and spatiotemporal objects (classes)	Spatial objects (classes)
Associations	Spatially, temporally and spatiotemporally dependent associations	Spatial, temporal and spatiotemporal associations	
Phases of system development	Analysis	Analysis Design Implementation	Design Implementation
Uses UML extension mechanisms	NO	YES	YES
Tool available	NO	YES	YES

Table 2: Summary of spatio-temporal modeling languages

When compared to ESTUML, it should be noted that although ST-PVL for UML is a conceptual visual modeling language, *it is a step further in the analysis process*. Also, it should be noted that ST-PVL for UML supports *a* *limited set of temporal* and consequently, spatiotemporal concepts (e.g. transaction time is not supported). On the other hand, having a tool that supports the modeling language and goes from analysis though design to implementation is an important strength of ST-PVL for UML. Finally, an important weakness of ESTUML is that no visual modeling tool exists, to support it.

The ArcInfo UML Geodatabase Model on the other hand, is a spatial model that may be used during design in case the implementation platform chosen is ArcInfo. Unfortunately, the model as well as the implementation platform does not support the time dimension; time modeling and implementation is left to the developer of the system.

The importance of spatial representation in residential mobility systems necessitates the use of a powerful GIS platform such as ArcInfo. In addition, ArcInfo exploits UML, a powerful analysis and design tool by incorporating the ArcInfo UML Geodatabase Model. Therefore, within the context of this work ArcInfo will be used for the implementation of *URM-Microsim* and therefore, the ArcInfo UML Geodatabase Model will be used during design. However, a conceptually rich model that supports time is required during analysis and the Extended SpatioTemporal UML (ESTUML) is chosen.

3.10 The STUML Process

Since neither ESTUML nor ArcInfo UML Geodatabase Model defines a process, one should be provided to guide the modeler during the construction of the system model. The process that will be used is called SpatioTemporal UML (STUML).

The STUML process is a set of activities that leads to the orderly construction of the system under development by applying the UML modeling techniques extended with spatiotemporal constructs. It is based on the TUML process (Svinterikou, 1998). The TUML process develops the Booch process (Booch, 1994) to address temporal concerns by:

- Modifying the Booch process to incorporate UML modeling techniques. In particular, the TUML process further modifies the Booch process to exploit the use of the use case diagram technique and to provide support for requirements capture. As the rest of the UML modeling techniques are mainly derived from the ones provided by the Booch method, they fit well in the steps and guidelines of the Booch process.
- Extending the Booch process with steps and guidelines that facilitate the capture of the temporal semantics of information.

The STUML process further develops the TUML process (Svinterikou, 1998) to address spatial and spatiotemporal concerns by extending the TUML process with steps and guidelines that facilitate the capture of the spatial and spatiotemporal semantics of information.

The STUML process is architecture-centric, iterative and incremental. In addition, it is *use-case driven*, that is, the whole development process is controlled by use cases and consequently by the user requirements. The process allows creativity and innovation within a controlled development process by distinguishing the micro and macro elements of the development process. The micro process serves as the framework for an iterative and incremental approach to development while the macro process serves as the controlling framework for the micro process.

During the macro process a set of activities takes place that leads to a number of products and allows the development team to identify risks and make early corrections to the micro process. It represents the activities of the development team during the entire development life cycle. As Figure 19 shows, the high-level activities that comprise the macro process are:

- *Conceptualization*: establish the core requirements for the software.
- *Analysis*: develop a model of the system's desired behavior.
- *Design*: create an architecture for the implementation.
- *Evolution*: evolve the implementation though successive refinement.
- *Maintenance*: manage post-delivery evolution.

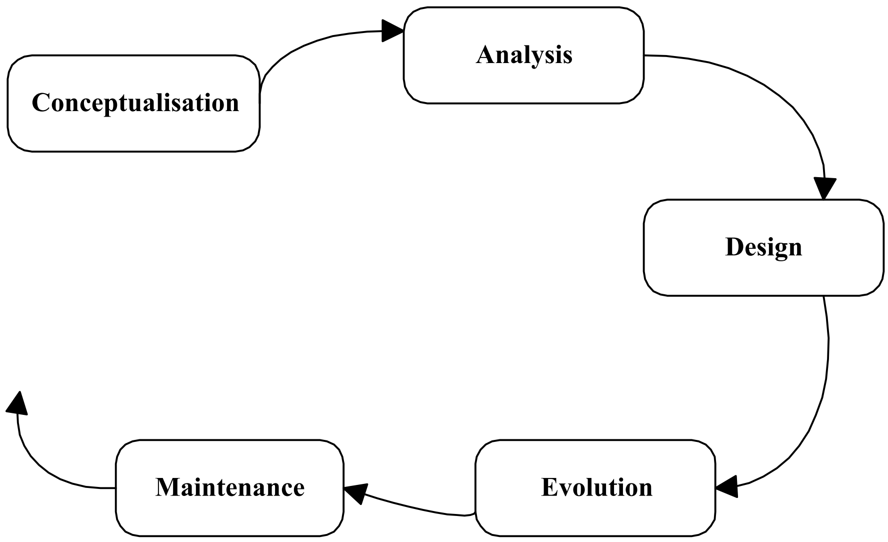


Figure 19: The Macro development process

The macro process repeats itself after major product releases. This particularly applies to the development of families of products.

The micro process is driven by the set of use cases and product releases that result from, and are successively refined by, the macro process. It mainly represents the daily activities of the individual developer or a small team of developers.

In the micro process the traditional phases of analysis and design are intentionally blurred. As Figure 20 shows, the activities that comprise the micro process are:

- *Identify the classes and objects at a given level of abstraction.*
- *Identify the semantics of these classes and objects.*
- *Identify the relationships among these classes and objects.*
- *Identify the spatial semantics of these classes, objects and relationships.*
- *Identify the temporal semantics of these classes, objects and relationships.*
- *Specify the interface and then the implementation of these classes and objects.*

.

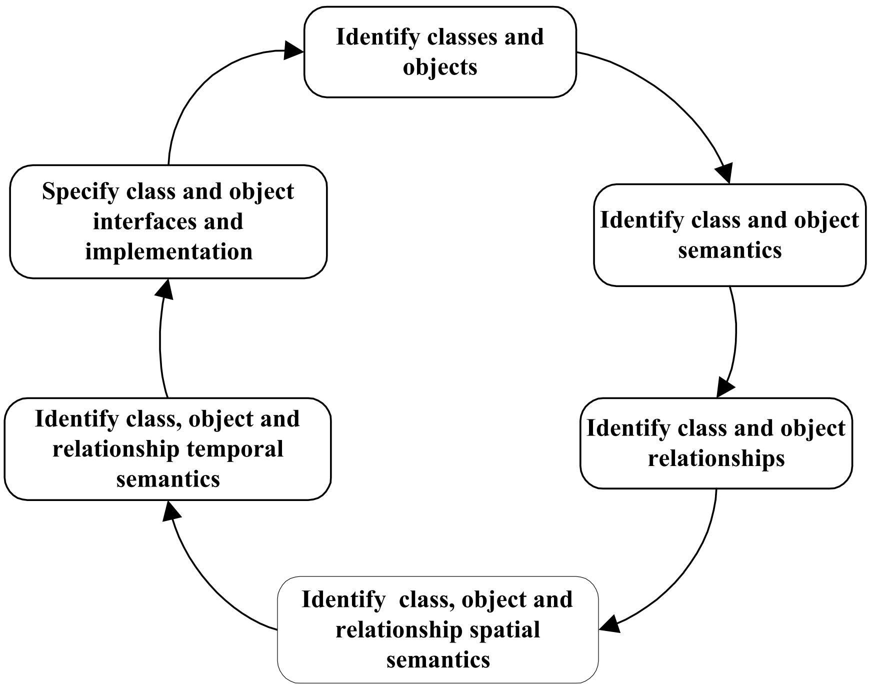


Figure 20: The Micro development process

The whole STUML process is presented in detail in **Appendix B**.

3.11 Summary

In this chapter, powerful computer science tools were discussed. Initially, the kinds of software system development methods were presented where the predominance of object-oriented technology was pointed out. Next, the elements of object-oriented methods were presented. Then, the Unified Modeling Language (UML) was presented, which is defined by the Object Management Group (OMG), as *the standard object-oriented* modeling language and the use of UML in microsimulation model development was illustrated, through an example: applying UML to the existing model LOCSIM. Next, the spatial, temporal and spatiotemporal concepts of systems were defined. In addition, two existing Spatio-Temporal modeling languages based on UML were discussed called Extended SpatioTemporal UML and SpatioTemporal PVL for UML, as well as the ArcInfo Geodatabase Model, and their main strengths and weaknesses were pointed out. Moreover, an overview of the modeling languages presented was given and the choices made within the context of this work were justified. Finally, ST-UML process was defined, that is, a set of activities that guides the development of a system, by applying the UML modeling techniques extended with spatiotemporal constructs.

Chapter 4
System Analysis

4.1 Introduction

In this chapter, the analysis model of *URM-Microsim,* i.e. the Housing Market System to be developed within the context of this thesis is presented. The analysis model captures the requirements of the system to be developed and does not contain any implementation details in order to be usable by all members of a development team (users, analysts, programmers). For the graphical representation of the analysis model, UML with spatio-temporal extensions is used in conjunction with the Microsoft Visio 2000 UML CASE tool. Sections 4.2 to 4.5 present the behavior of the system to be developed (what the system does) that is, the *URM-Microsim Model* as well as its components, namely, *the Housing Demand, Housing Supply* and *Residential Search and Migration* sub-models. Finally, section 4.6 presents the *URM-Microsim* Database, namely, the *entities* that are of interest as well as *their characteristics* and the *relationships* with each other.

4.2 The *URM-Microsim* Model

The *URM-Microsim Model* is a dynamic, spatio-temporal microsimulation model that captures the developments in demographic structures and housing markets, as well as the interrelation between them.

The model consists of three sub-models (*packages*): *Housing Demand*, *Housing Supply*, and *Residential Search and Migration* (Figure 21). The *Housing demand* sub-model simulates the occurrence of demographic events that influence the household decision to move such as death, fertility, marriage, and income change and therefore determines the number of households that will search for a new house. The *housing supply* sub-model on the other hand, simulates the changes in housing stock and thus determines the number of houses that are available to households. Finally, the *Residential Search and Migration* sub-model simulates the clearing of the housing market, that is, which households actually find a new dwelling. The outcome of the *Residential Search and Migration* sub-model affects the way housing demand and supply are modified.

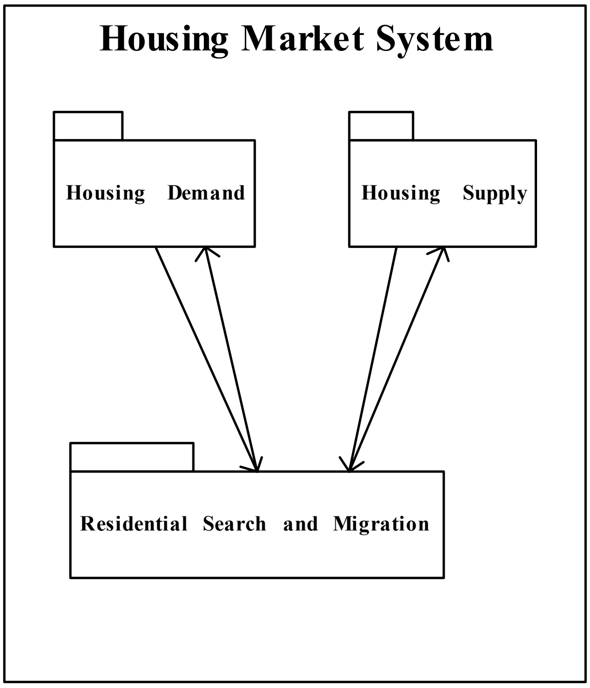


Figure 21: The URM-Microsim analysis model

4.3 The Housing Demand sub-model

Housing demand depends on several factors, such as changes in household structure caused by *demographic events* (e.g. mortality, fertility, marriage/divorce etc), *migration* (in- or out- migration), changes in *income and employment* of the urban population, and changes in *demand for* *specific housing services* (e.g. accessibility to employment location, structure type, neighborhood quality). As individuals and households pass through different life stages (fertility, marriage, occupation, education), their *housing preferences change*, triggering mobility (Oskamp, 1997). For example, a change of job (occupational life stage) could lead to an increase in income, which makes an upward move in the housing market possible. If the new job is not within acceptable commuting distance, it could also imply that a move must be made. All households do not have an equally strong desire to find a new dwelling; the *search intensity* with which a household searches for a new dwelling depends on the event that triggers the desire to move. For example, a divorcee is urged to find a new dwelling, while a household with a new child searches the housing market with low search intensity. Three types of move by destination are distinguished: households moving within the boundaries of the housing market (*residential mobility*), households moving from the housing market to another one (*out-migration*), and households moving into the housing market from the rest of the world (*in-migration*). Figure 22 shows the *use case diagram* corresponding to the *Housing Demand* sub-model. The sub-model affects *Individuals* as well as *Households.* In addition, it models three types of events: *Demographic change*, *Income change* and *Demand change* events. The *use-cases scenarios* that define these events in detail are presented in **Appendix C.**

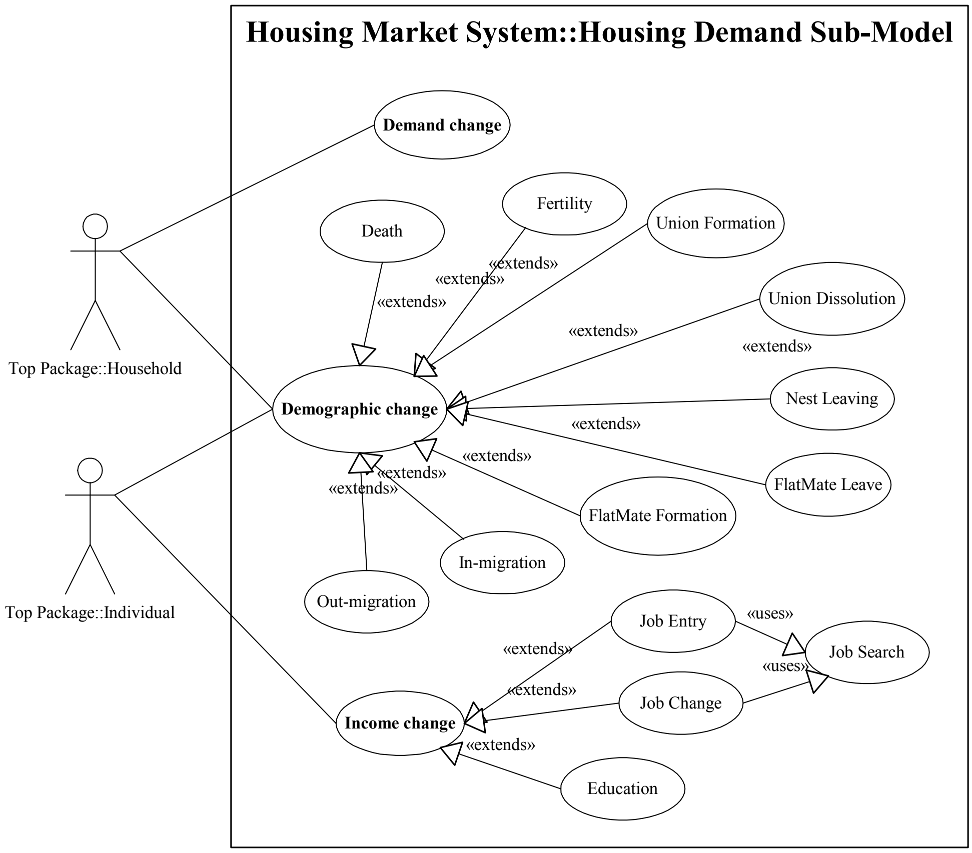


Figure 22: The Housing Demand analysis sub-model

4.3.1 Demographic change events

Each individual is subjected to the probability of experiencing one of the *death*, *fertility, union formation, union dissolution, flat-mate formation, flat-mate leaving, nest leaving* events and *out-migration*, based on the individual's current life cycle stage, as well as the previous event experienced by the individual's household. In addition, the number of *in-migrating* households is determined.

- *Death.* Each individual is subjected to a probability of dying based on his/her age and sex. If the individual dies the information concerning the individual is updated. In addition his/her household's properties are updated as well as the information concerning the remaining household members. The household expresses the intention to move to another dwelling.

- *Fertility.* Each female individual is subjected to a probability of giving birth based on her age and marital status. If the individual gives birth to a child, the information concerning the child is determined and the mother's household properties are updated. The household expresses the intention to move to another dwelling.
- *Union formation.* Each not married individual of certain age searches for a not married individual of the opposite sex and of certain age. If the individual finds a partner, then the couple decides to live together (form a household) and to search for another dwelling.
- *Union dissolution.* Each couple (married or cohabitates) is subjected to a probability of splitting based on their age and the years of marriage. If the couple splits, it is decided which partner leaves the house. The partner that leaves the house forms a new household and decides to search for another dwelling.
- *Nest leaving.* Each individual of age 18 and above, who still lives with his/her parents is subjected to a probability of moving out. If the individual decides to leave the parental house, he/she forms a new household that must search for another dwelling.
- *Flat-mate formation.* An individual who lives on his/her own or flat-mates with others is subjected to the probability of finding another flat-mate. If the individual finds a flat-mate, then they decide to flat-mate (form a household) and to search for another dwelling.
- *Flat-mate leaving.* An individual who flat-mates with other individuals is subjected to the probability of leaving the household to live on his/her own. If the individual decides to leave the initial house, he/she forms a new household that must search for another dwelling.

- *Out-migration.* An individual or a household may decide to out-migrate due to several reasons. For example, an individual who is unemployed for a long time may decide to out-migrate. A child of age 18 and above may decide to study at the University of another city. If the whole household out-migrates, then its dwelling is vacated and added to the housing supply set.
- *In-migration.* Households outside the housing market decide to in-migrate. These households consist of students, army officers, foreign workers etc. These households must find a dwelling.

4.3.2 Income change events

Each individual of age 18 and above is subjected to the probability of having his/her income changed based on the individual's education career and occupational changes.

- *Education.* The education group to which the individual belongs is determined (none, primary or less, lower secondary, upper secondary, post secondary, tertiary) and he/she is assigned an education career.
- *Job entry.* Each individual of age 15 and above who does not work (soldiers are excluded) is searching for a job according to the first criteria of the job search event (the remaining criteria are not taken into consideration). If the individual finds a job, then his/her household expresses the intention to move to another dwelling.
- *Job change.* Each individual, who already works, is subjected to a probability of retiring or becoming unemployed. Alternatively, the individual searches for a new job according to the criteria of the job search event. If a job change takes place, then the individual's household expresses the intention to move to another dwelling.
- *Job search.* The individual searches for a job according to the following criteria:

1. The type of occupation should correspond to the education group to which the individual belongs.
2. The job offers a higher salary than the one offered by the worker's current job based on the worker's education group and the type of occupation.
3. The job is located at the city center or at the residence zone of the individual's house or neighboring residence zones.

If the individual finds a job that matches his/her criteria, then he/she accepts the job offer.

4.3.3 Demand change event

A household that does not have experienced a demographic or income change event is subjected to the probability of searching for a new dwelling due to demand changing.

- A household whose rent and housing expenditure are lower than the household's housing budget is simulated to request housing in a more attractive neighborhood aiming for socio-economic upgrading.
- A household whose rent and housing expenditure exceed the household's housing budget is simulated to request a less expensive dwelling.

4.4 The Housing Supply sub-model

Housing supply develops on the long-term to satisfy housing demand. Changes on housing supply arise from the construction of *new dwellings*, the *demolition* and *conversion* of existing dwellings (e.g. splitting, combining and renovation) and the *changes in the housing expenditures* of dwellings. Figure 23 shows the *use case diagram* corresponding to the *Housing Supply* sub-model. The sub-model affects *Dwelling* and models four types of events: *New construction*, *Structure conversion*, *Demolition* and *Housing expenditures change*. The *use-cases scenarios* that define these events in detail are presented in **Appendix C.**

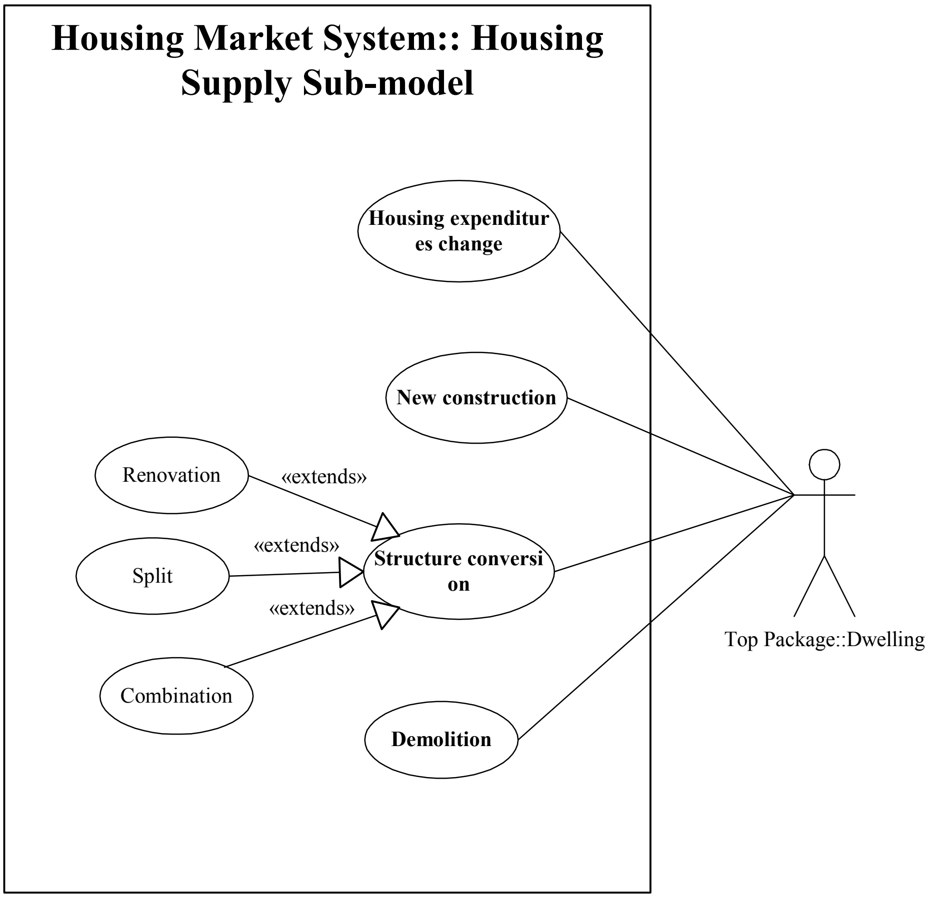


Figure 23: The Housing Supply analysis sub-model

4.4.1 New construction event

The number of new dwellings to be added to the house stock is determined based on the market's mobility and housing shortage. The location of each new dwelling is determined based on the undeveloped land suitable for development of the market area and the market's housing demand trends. The dwelling's type and services are determined based on the market's housing demand trends.

4.4.2 Structure conversion events

The structure of a dwelling is converted according to some possible way.

- *Split.* Each vacant, large dwelling is subjected to the probability of splitting to two or more smaller dwellings. The cost and difficulty of the split are determined based on how similar the current and targeted structure types and construction qualities are. If the split is profitable, then the split takes place.
- *Combination.* All set of two or more vacant, small and neighboring dwellings are subjected to the probability of being combined to a larger dwelling. The cost and difficulty of the combination are determined based on how similar the current and targeted structure types and construction qualities are. If the combination is profitable, then the combination takes place.
- *Renovation*. Each dwelling in medium or poor condition is subjected to the probability of being renovated. The cost and difficulty of the renovation are determined based on the current and targeted construction qualities of the dwelling. If the renovation takes place, then the type of renovation is determined (dwelling maintenance, dwelling extension, number of rooms change etc).

4.4.3 Demolition event

A dwelling in poor condition is subjected to the probability of being demolished. If the dwelling is occupied, then the corresponding household must search for another dwelling. The demolition takes place only after the household moves out.

4.4.4 Housing expenditures change event

The housing expenditures (maintenance capital, heat, electricity etc) of each dwelling is determined based on the dwelling's type, size, age and construction quality.

4.5 The Residential Search and Migration sub-model

Households that are dissatisfied with their current housing, *search* the *housing market* for a vacant dwelling that meets their needs. If a household finds a dwelling that complies with its preferences, then the household *moves* into the new dwelling. If household search fails, then the household either *remains* to its current dwelling or migrates to another housing market (*out-migration*). Housing demand and supply determine the rent and market value of the dwellings (*housing prices*).

Figure 24 shows the *use case diagram* corresponding to the *Residential Search and Migration* sub-model. The sub-model affects *Households* as well as *Dwellings* and models *Search and Migration* and *Housing prices* events. The *use-cases scenarios* that define these events in detail are presented in **Appendix C.**

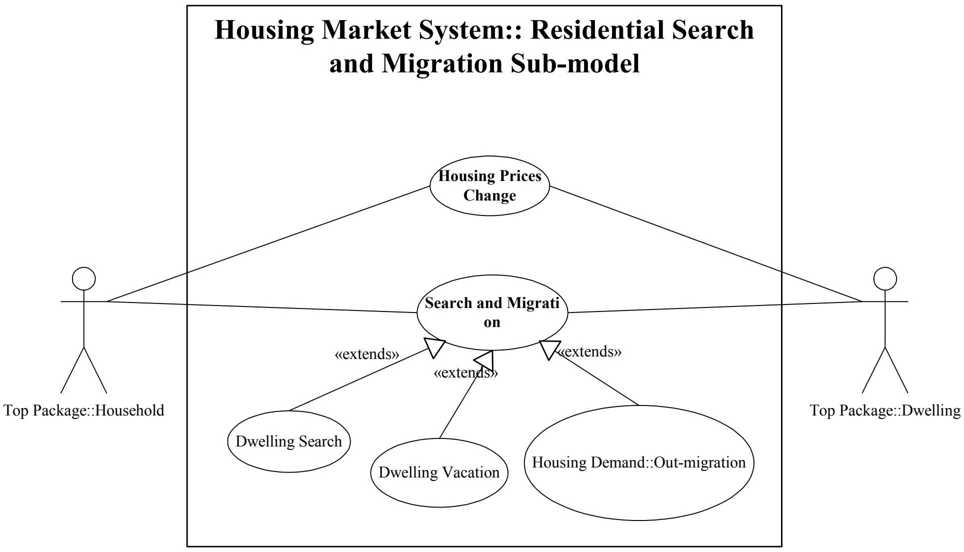


Figure 24: The Residential Search and Migration analysis sub-model

4.5.1 Search and migration event

All households whose structure or housing circumstances have changed and all newly formed households as well as the households they originate from are simulated to search for a new dwelling.

- *Dwelling Search.* The household searches the list of vacant dwellings for a dwelling according to the following rules:
    - *Tenure*. Initially, the household decides whether it will search for a house to own or rent, based on the household's *previous tenure*, its current *income*, its current *life cycle stage* and the current *mortgage interest rates* and *taxation allowances* (e.g. households that decide to buy a dwelling and do not already own one are entitled to a tax relief).
    - Dwellings are defined by seven (7) dimensions: *dwelling price/rent, number of rooms, dwelling type, dwelling age*, *structure condition, amenities (garden, parking etc),* and *neighborhood quality.* The household defines the values of the dimensions the desired dwelling should have, as well as the dimensions that are of importance

(primary dimensions) and those that are not (secondary dimensions). The desired dwelling is defined based on the event experienced by the household, the household's income as well as whether the household searches for a dwelling to rent or own.

  - The *number of dwellings considered* by the household depends on the *search intensity* of the household. In other words, households with high search intensity consider more dwellings than the ones with low search intensity.

- If the household decides to accept a vacancy then it moves to the new dwelling and the dwelling the household occupied (if any) is vacated (*Dwelling vacation).*
- If the household fails to find a dwelling, then if it is an in-migrating household, the household as well as its members is deleted. In any other case, the household is subjected to the probability of *out-migrating*.
- If the household does not out-migrate then if it already has a dwelling it remains there, or in any other case the household is deleted and its members return to their parental households.

4.5.2 Housing prices change event

Once all household moves have taken place, the rent and market values of the dwellings are updated based on housing demand and supply.

4.6 The *URM-Microsim* Database

Within *URM-Microsim*, the database entities of interest are *individuals*, *households* and *dwellings*. The population of a city consists of *households*. A household has one or more members. Each household member is an *individual* with a specific position in the household (couple member, child or single) which may change over time. An individual may have an *occupation* (and therefore, a salary) or not which may change over time. The occupation belongs to a *residence zone*. A household lives at a *dwelling* that owns or rents and that may change over time. A dwelling may have at some point in time one or more ancestors (the dwelling is the outcome of a split or combination) or one or more descendants (the dwelling takes part in a combination or split process). A dwelling belongs to a *building*, which in turn belongs to a *building block*, which in turn, belongs to a *residence zone*. The residence zone has one or more neighbours. Buildings, building blocks and residence zones have a position in space i.e. they are spatial objects. In addition, the database entities have characteristics, whose values may change over time as well as the relationships between the entities.

An *individual* has the following characteristics:

- Personal characteristics
  - *Sex.*
  - *Date of birth.*
  - *Marital status* (never married, married, divorced etc).
  - *Date of in-migration.*
  - *Date of out-migration.*
  - *Date of death.*

  The marital status of an individual may change over time and this change should be kept in the database. That is why marital status should be denoted as a temporal attribute. Moreover, individuals who out-migrate or die are not deleted from the database in order to store the past of the housing market system. That is why an individual should be denoted as a temporal object.
- Occupational characteristics
  - The *education* group to which the individual belongs (none, primary or less, lower secondary, upper secondary, post secondary, tertiary).

- The *employment* status of the individual (employee, employer, self-employed, family worker, unemployed, student, engaged in family duties, engaged in the army service, retired, other inactive).
- The *search intensity* with which the individual searches for a new job (none, low, high).
- The *income* of the individual which is calculated as the sum of the individual' *occupational income*, that is:
  - *Salary* if the individual is a worker.
  - *Unemployment allowance* if the individual is unemployed.
  - *Pension* if the individual is retired.
  - The individual's *non-occupational income* (e.g. rent of a house owned by the individual etc).

  All occupational characteristics of an individual may change over time. The education level, employment status and income changes are of interest within the context of this work and should be kept in the database. That is why these characteristics should be denoted as temporal attributes.

An *occupation* has the following characteristics:

- *Occupation branch* (professional, technician and associate professional, clerk, service worker and shop and market sales worker, etc).
- The *industry branch,* that is, the branch of economic activity to which the occupation belongs (construction, wholesale and retail trade, hotels and restaurants, public administration and defence, education, etc).
- *Working time* (full-time, part-time).

A *household* has the following characteristics:

- The *household size* that is, the number of individuals that constitute the household.

- The *event* that the household currently experiences.
- The *search intensity* with which the individual searches for a new house (none, low, medium, high).
- The household's *income*, which is calculated as the sum of the household members' income.

All household characteristics may change over time. Since household size changes and all the events experienced by the household during its life cycle, are of interest in order to capture population life histories, they should be denoted as temporal attributes. In addition, households which no longer belong to the housing market under examination are not deleted from the database in order to store the past of the housing market system. That is why a household should be denoted as a temporal object.

A *dwelling* has the following characteristics:

- *Dwelling Size.*
- *Number of rooms*.
- The dwelling's *structure quality*[4] (quality of interior and exterior finishing, the architectural distinction of the unit, and the unit's spaciousness or average room size).
- *Renovation year*.
- *Renovation type,* that is an indicator to denote the type of renovation that took place if any (room add/change, structure renovation, building renovation).
- *Market value* of the dwelling.
- *Rent value* of the dwelling.
- *Housing expenditures* of the dwelling.

[4] In [HUDS] it is called *Quantity of structure services.*

- *Available*, that is an indicator to denote whether the dwelling is available to searching households or not (no, for rent, for sale).

All dwelling characteristics may change over time and their changes should be kept in the database. Therefore, all of them are denoted as temporal attributes. Moreover, demolished dwellings are not deleted from the database in order to store the past of the housing market system. That is why a dwelling should be denoted as a temporal object.

A *building* has the following characteristics:

- *Building type* (one, two, three or more dwelling house).
- The *size of the lot* to which the building belongs.
- *Number of floors*
- *Number of dwellings*
- *Parking*, that is, an indicator to denote whether the building offers parking place or not.
- *Garden*, that is, an indicator to denote whether the building has a garden or not.
- The *construction period* of the building (age of the building).
- *Demolition year.*

The type, number of floors and number of dwellings of a building may change over time due to its dwellings' renovation. Since these changes are of importance, the aforementioned attributes should be denoted as temporal ones.

A *building block* has the following characteristics:

- The level of *shopping access* of the neighborhood.
- The level of *school access* of the neighborhood.
- The level of *transportation access* of the neighborhood.

A *residence zone* has the following characteristics:

- *Zone name*.
- *Zone type*.
- *Zone quality*.
- *Population type*.

Figure 25 shows the *Class diagram* corresponding to the *URM-Microsim*. This diagram shows the database structure of the system (static structure of the system), that is, the *classes*, their *attributes*, *operations*, and *relationships* with other classes that will be implemented and maintained in the database of the Housing Market System to be developed.

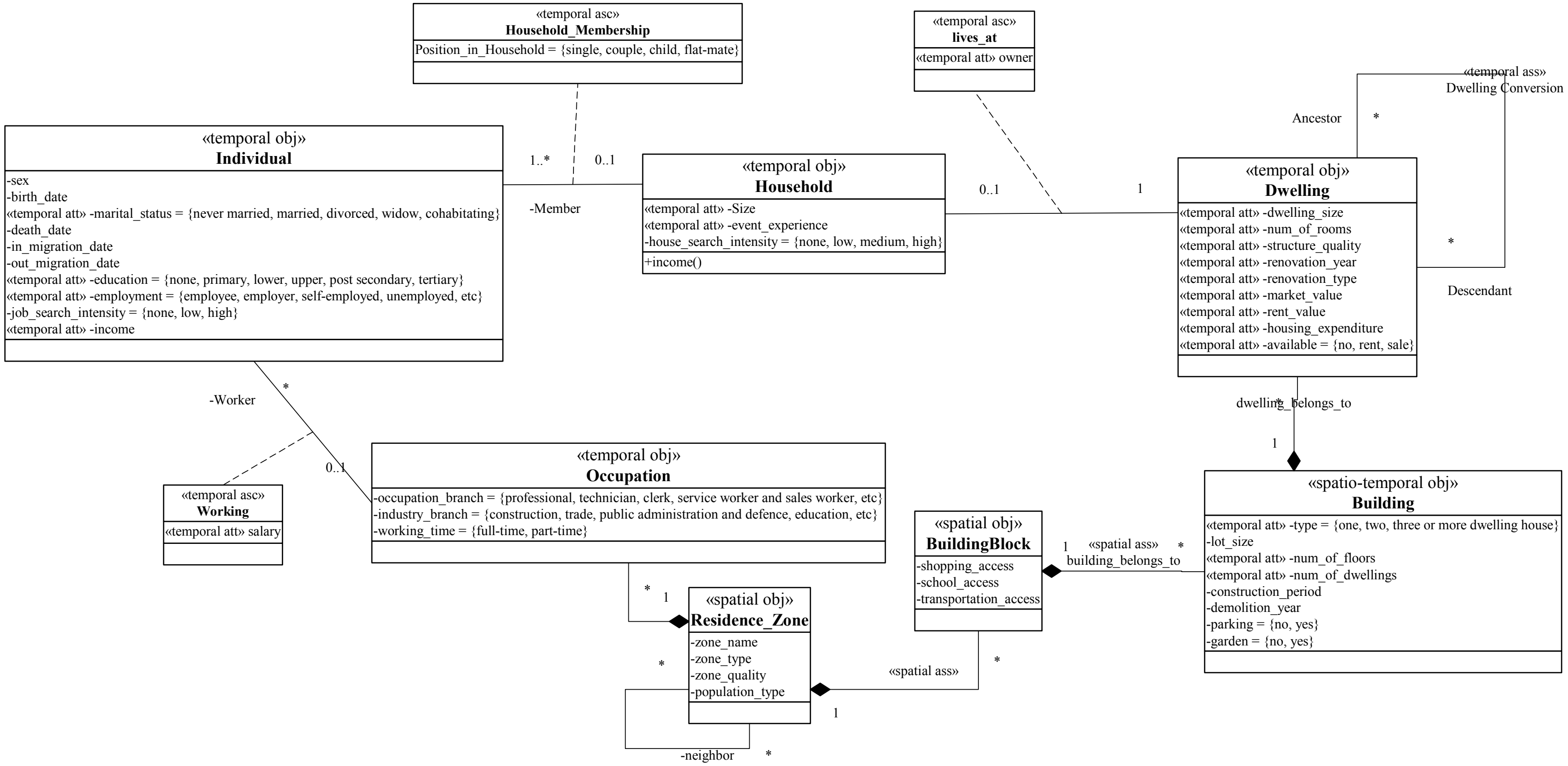


Figure 25: The URM-Microsim Database – Class Diagram

4.7 Summary

In this chapter, the analysis model of Housing Market System to be developed within the context of this thesis was presented. UML with spatio-temporal extensions was used for the graphical representation of the analysis model. Initially, the behavior of the system to be developed (what the system does) was presented that is, the *URM-Microsim Model* as well as its components, namely, *the Housing Demand, Housing Supply* and *Residential Search and Migration* sub-models. Then, the *URM-Microsim* Database was presented, namely, the *entities* that are of interest as well as their *characteristics* and the *relationships* with each other.

Chapter 5
System Design

5.1 Introduction

In this chapter, the design model of *URM-Microsim* is presented taking into consideration the implementation platform to be used. Therefore, UML with ArcInfo extensions (ArcInfo Geodatabase Model) in conjunction with the Microsoft Visio 2000 UML CASE tool is used for the graphical representation of the design model. Section 5.2 presents the behavior of the *URM-Microsim Model* and its sub-models *(Housing Demand, Housing Supply* and *Residential Search and Migration)*, that is, the steps the model goes through, in order to execute the model processes. Section 5.3 presents the events that occur in the Housing Demand sub-model. Section 5.4 presents the events that occur in the Residential Search and Migration sub-model. Additionally, section 5.5 presents the *URM-Microsim* Database, namely, the *entities* that are of interest as well as *their characteristics* and the *relationships* with each other. Finally, section 5.6 presents implementation issues concerning the implementation of the *URM-Microsim* design model.

5.2 The URM-Microsim Model

The *URM-Microsim Model* and its sub-models *(Housing Demand, Housing Supply* and *Residential Search and Migration)* defined in the analysis model are further elaborated to show the steps the model goes through, in order to execute the model processes.

In *URM-Microsim* the housing market under examination is initially determined, namely the city whose residential mobility will be simulated. In addition, the date of simulation (execution date) is defined. Then the *Housing Demand* and *Housing Supply* sub-models of *URM-Microsim* are executed in parallel.

In the *Housing Demand* sub-model each individual of the population is initially subjected to the probability of experiencing one of the *income* events (*education, job entry* and *job change*) and successively is subjected to the probability of experiencing one of the *demographic* events (*death*, *fertility, union formation, union dissolution, flat-mate formation, flat-mate leaving,* and *nest leaving*). Next, each household that does not have experienced a demographic or income change event is subjected to the probability of experiencing the *demand change* event. In addition, the type and number of in-migrating households is determined. Each in-migrating household is simulated to experience the *in-migrating* event.

In the *Housing Supply* sub-model each dwelling of the housing stock is initially subjected to the probability of experiencing one of the *structure* events (*renovation, split, combination* and *demolition*) and successively is subjected to the probability of experiencing the *housing expenditure changes* events. In addition, the type and number of new dwellings is determined. Each new dwelling is simulated to experience the *new construction* event.

Once *Housing Demand* and *Housing Supply* sub-models' execution is completed, the *Residential Search and Migration* sub-model is initiated. All Households that wish to search for a dwelling are selected as well as all vacant dwellings. Each one of these households is randomly selected and the number of dwellings it will be offered is determined based on the household's *search intensity*. Next the *dwelling search* event is simulated. If the household fails to find a dwelling, then if it is an in-migrating household, the household as well as its members is deleted. In any other case, the household is subjected to the probability of *out-migrating*. If the household does not out-migrate then if it already has a dwelling it remains there, or in any other case the household is deleted and it's members return to their parental households. Once all households have searched the housing market and all household moves have taken place, the *housing prices* are updated based on housing demand and supply.

Figure 26 shows the *activity diagram* corresponding to the *URM-Microsim* design model under development. This diagram shows the steps the model goes through, in order to execute the model processes and explores how various processes interact in the model.

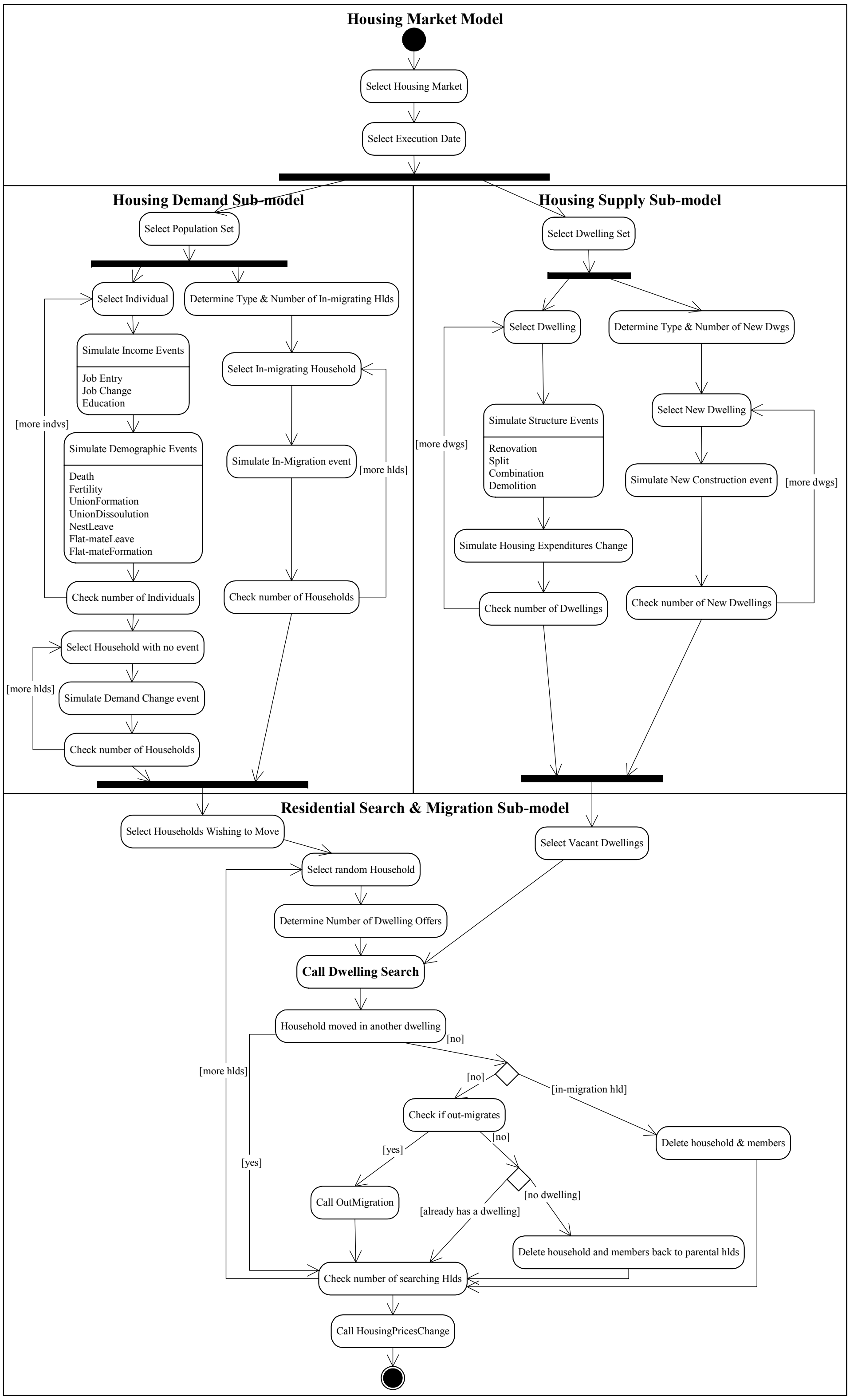


Figure 26: The URM-Microsim design model

5.3 The Housing Demand events

In this section each demographic and income event of the *Housing Demand* sub-model (Figure 26) is further elaborated.

5.3.1 Death event

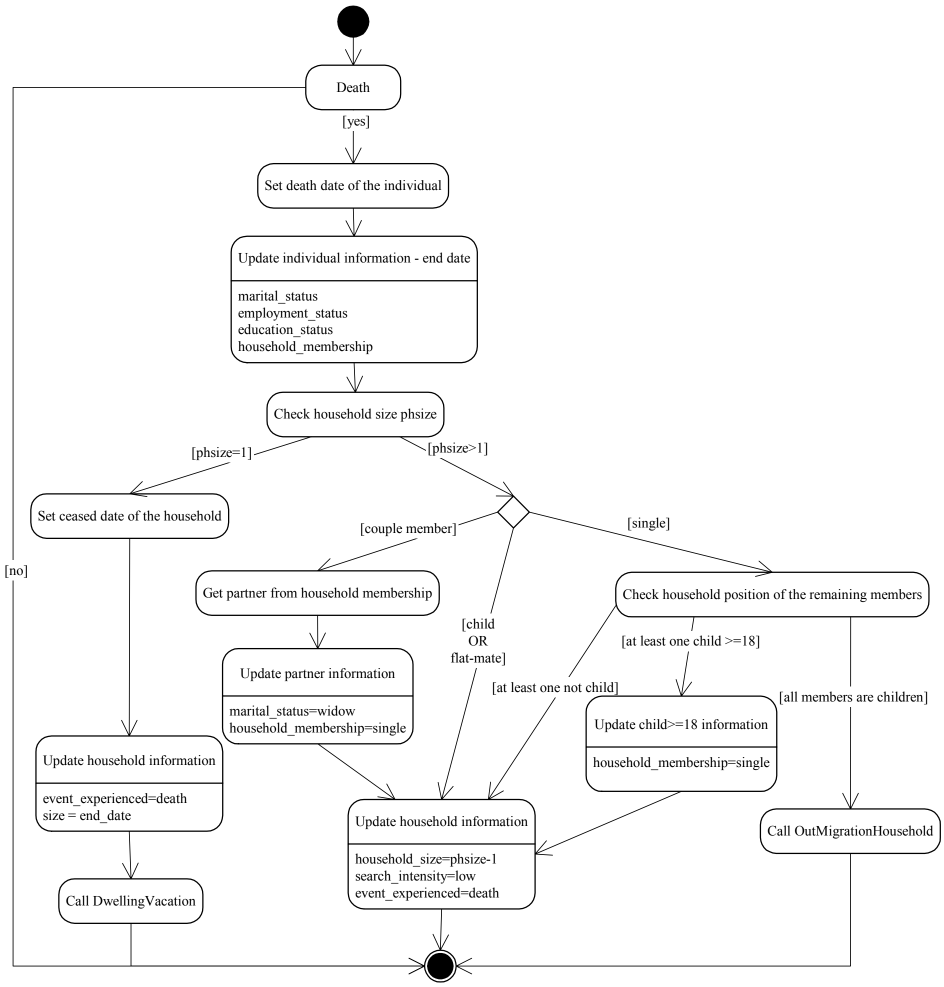


Figure 27: Death event

Each individual is subjected to a probability of dying based on his/her age and sex. If the individual is simulated to die then the death date of the individual is set and his/her information is updated to denote that this individual no longer belongs to the population. In addition, the household's information is updated and it is denoted as a household with *low search* *intensity*. If the individual belonged to a one-member household, then the household's ceased date is set to denote that the household no longer belongs to the city’s household set and the dwelling the household occupied is vacated and added to the housing stock. If the individual had a partner (he/she was a couple member) then the partner’s information is updated. If the individual was the parent in a one-parent household with one or more children, then the definition of a guardian (single or child>=18 years old) from the remaining household members is considered. If a guardian is available then his/her information is updated, otherwise the household out-migrates since it is assumed that the city does not have an orphan house.

5.3.2 Fertility event

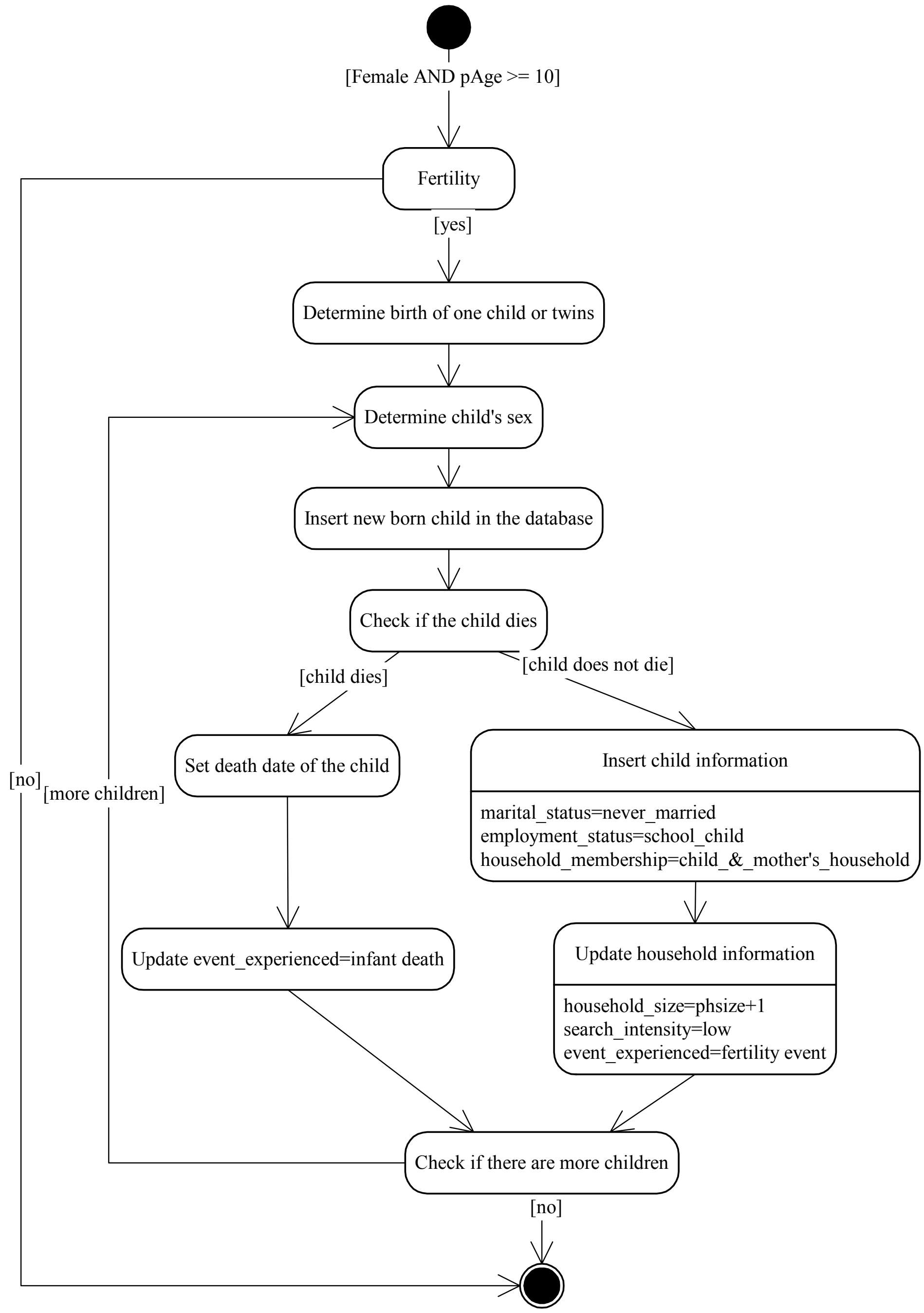


Figure 28: Fertility event

Each woman of age 10 and above is subjected to a probability of giving birth based on her age and marital status. If the woman is simulated to give birth then she is subjected to the probability of giving birth to twins. For each child born the child's sex is determined and he/she is added to the population. Next, each child is exposed to the probability of dying (infant mortality). If the child dies then the death date of the child is set and the household's information is updated. If the child survives, then information concerning the child is inserted, the household's information is updated and it is denoted as a household with *low search intensity*.

5.3.3 Union Formation event

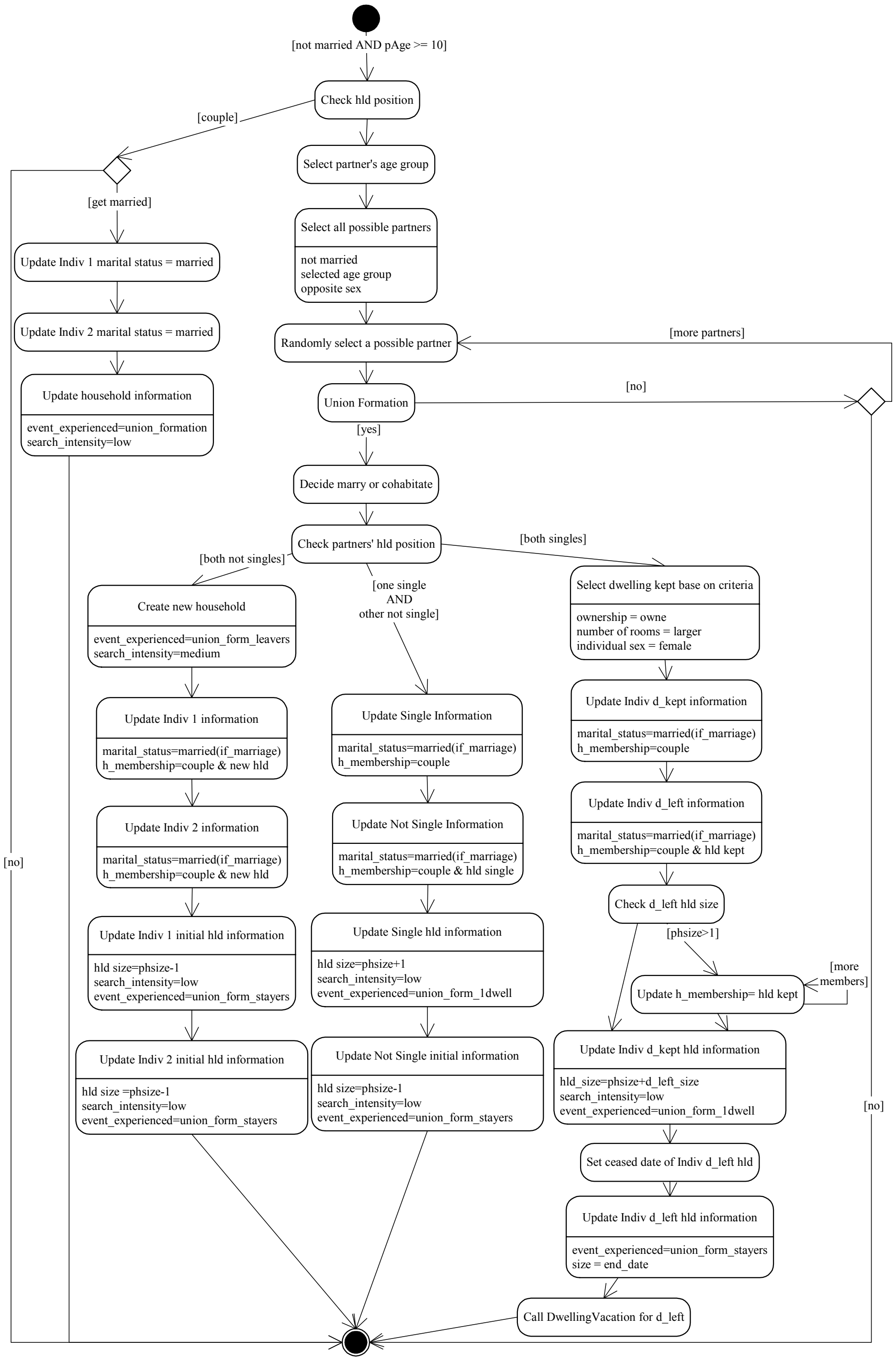


Figure 29: Union Formation event

Each individual of age 10 and above, who is not married, searches for a not married individual of the opposite sex and of certain age.

If the individual is a couple member (but not married) namely the individual cohabitates then he/she is subjected to the probability of getting married. If the individual marries his/her cohabiter then the information concerning the individual, the partner as well as their household is updated.

If the individual is not a couple member then the age group of possible partners is selected. Then the list of possible partners is selected, which consists of all not married individuals of the opposite sex and of the selected age group. A possible partner is randomly selected and the individual is subjected to the probability of accepting the partner based on his/her sex and age as well as, the partner's age. Partners are randomly selected until one is accepted or until all possible partners are examined.

If the individual finds a partner, then it is decided whether the individuals will get married or will cohabitate. Next the current household position of the two partners is examined. If both individuals are not singles that is, they either live with their parents or flat-mate, then a new household is created, which is denoted as a one with *medium search intensity*. The information concerning the partners as well as their parental households is updated. In addition, the parental households are denoted as households with *low search intensity.* It should be noted that the parental households will search for another dwelling only if the new household finds a new dwelling and consequently, the union formation actually takes place.

If only one of the partners is single that is, he/she already has a dwelling of his/her own, then the other partner moves in with him/her and become member of his/her household. The information concerning the partners, the single's household as well as the other partner's parental household is updated accordingly. In addition, both households are denoted as households with *low search intensity.*

If both partners are single that is they both have their own dwellings then it is decided who's dwelling and consequently household the partners keep. The two dwellings are compared based on ownership status, size (number of rooms) and partner's sex. Once it is decided whose partner dwelling is kept, the other partner as well as his/her children (if any) moves in with him/her and become member of his/her household. The information concerning the partners, the moving children (if any) as well as the household kept is updated accordingly. The household left has its ceased date set to denote that the household no longer belongs to the city's household set and the dwelling the household occupied is vacated and added to the housing stock. In addition, the household kept is denoted as a one with *low search intensity.*

5.3.4 Union Dissolution event

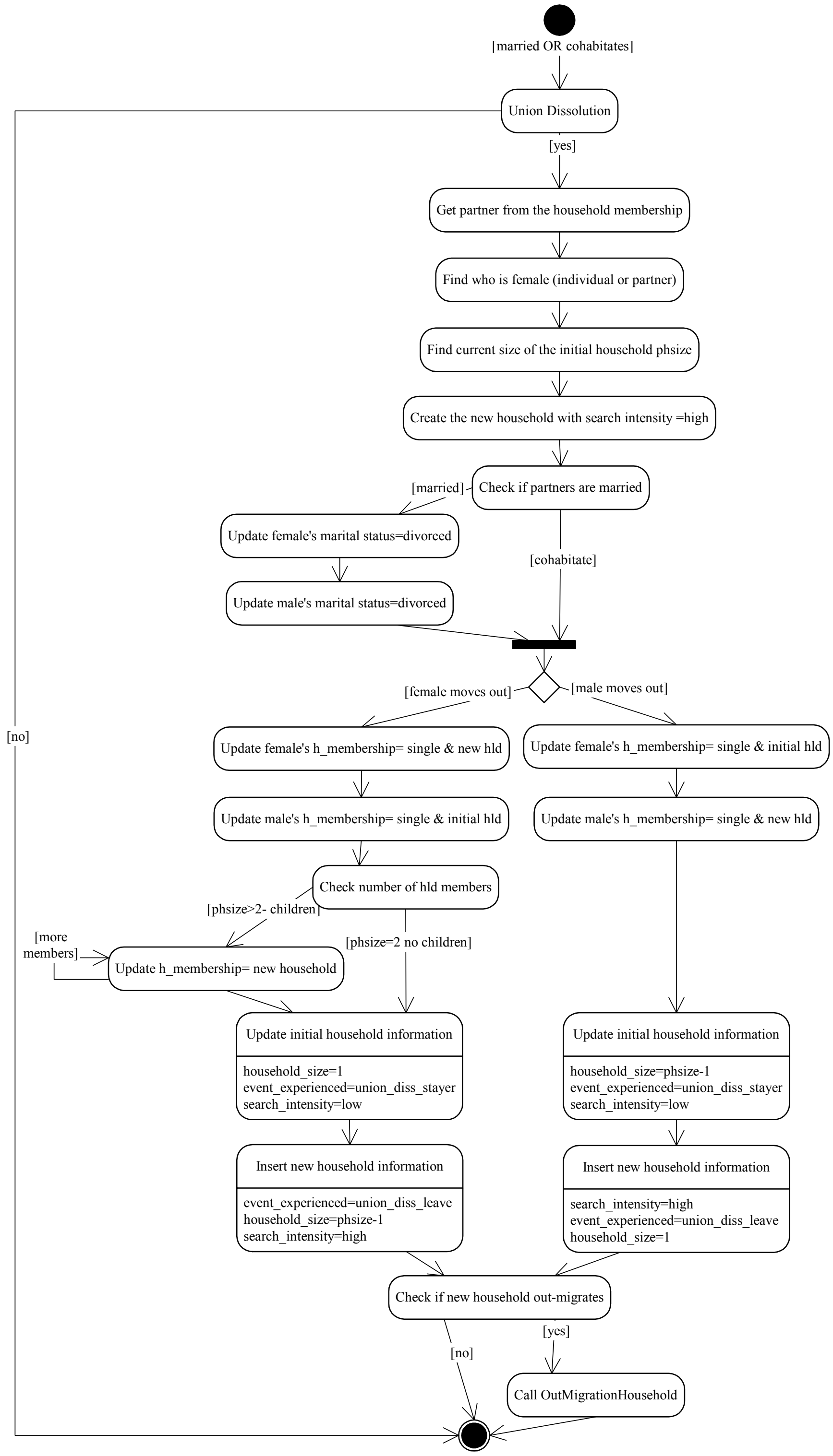


Figure 30: Union Dissolution event

Each individual who is married or cohabitates is subjected to the probability of separating based on female's age and the years of marriage/cohabitation. If the individual is simulated to separate from his/her partner, then it is decided which partner (male or female) moves out of the current dwelling. It is assumed that the children stay with the mother and any singles are the mother's parents. Next, a new household is created that consists of the partner that moves out and is denoted as a household with *high search intensity*. The information concerning the initial household is updated and the household is denoted as one with *low search intensity.* It should be noted that the initial household will search for another dwelling only if the new household finds a new dwelling and consequently, the union dissolution actually takes place. The partners' information is also updated as well as the information concerning the rest of the members, if any (in case the mother's household is moving out). Finally, the new household (moving out household) is subjected to the probability of out-migrating.

5.3.5 Nest Leaving event

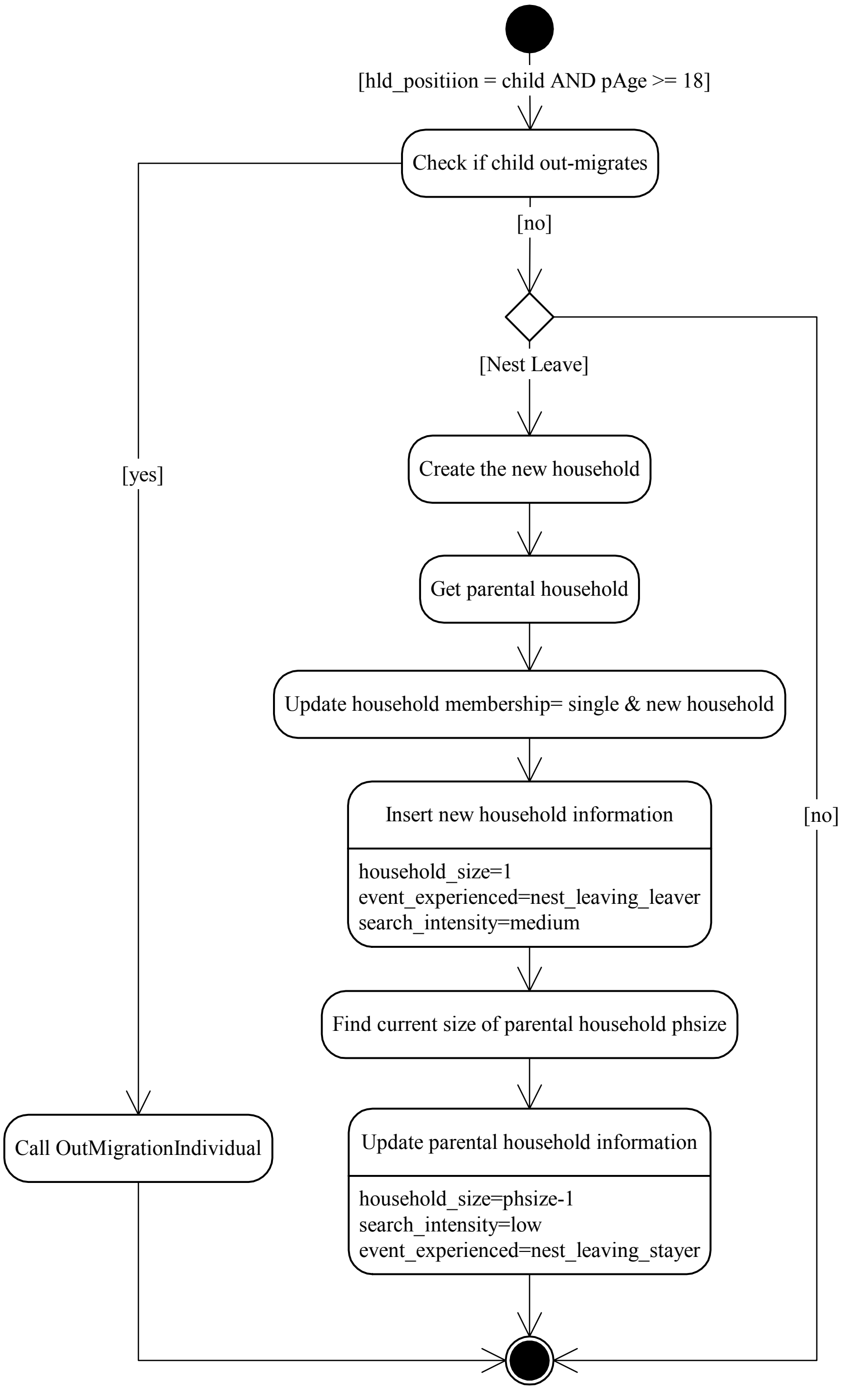


Figure 31: Nest Leaving event

Each individual of age 18 and above, who still lives with his/her parents is initially subjected to the probability of out-migrating. If the individual remains in the housing market system under examination then he/she is subjected to the probability of moving out the parental household. If the individual is simulated to leave his/her parents then he/she forms a new single-member household, which is denoted as a one with *medium search intensity*. The information concerning the individual as well as the parental household is updated. In addition, the parental household is denoted as one with *low search intensity.* It should be noted that the parental household will search for another dwelling only if the new household finds a new dwelling and consequently, the nest leave actually takes place.

5.3.6 Flat-mate Formation event

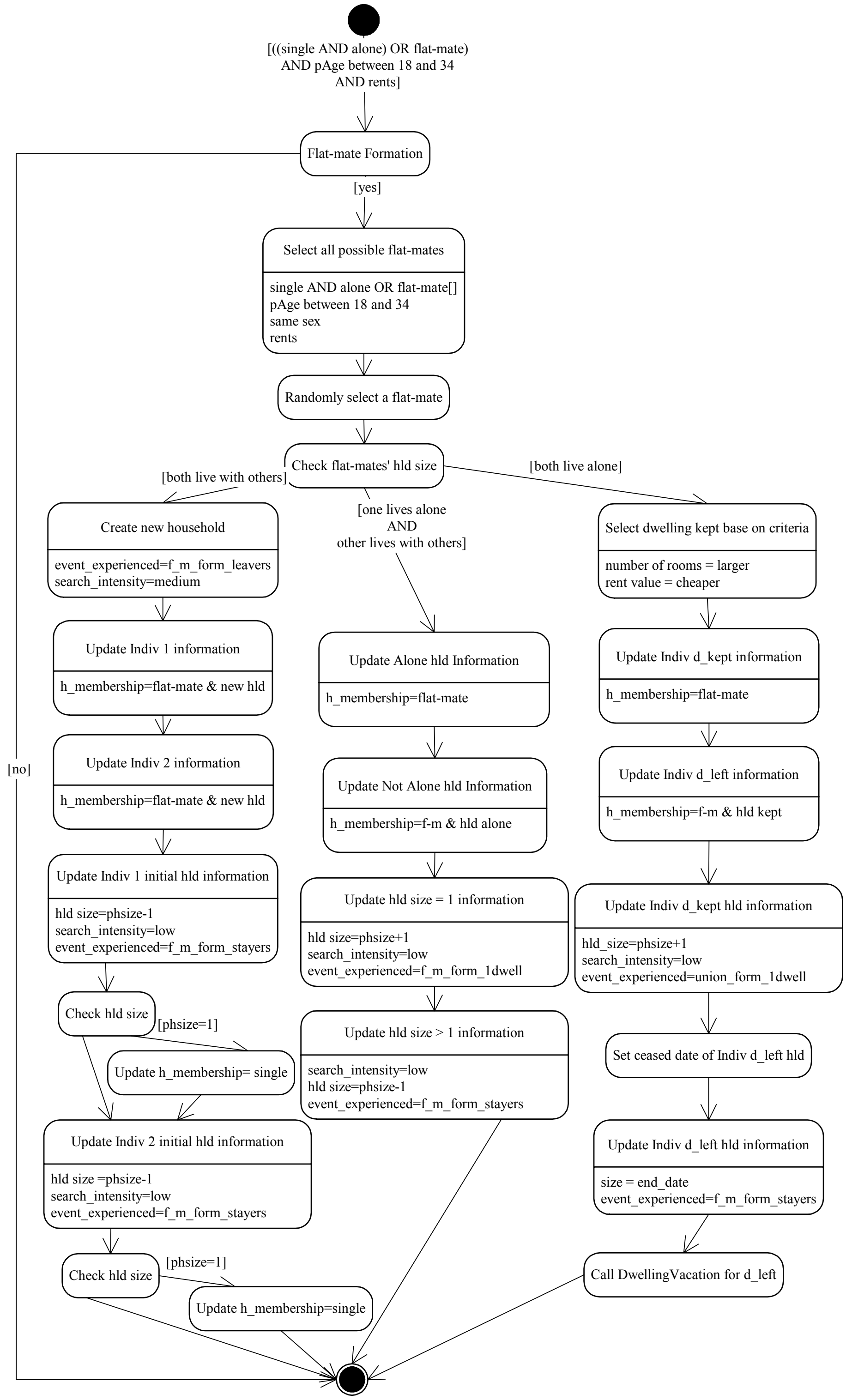


Figure 32: Flat-mate Formation event

Each individual of age between 18 and 34 who lives on his/her own or flat-mates with others in a rented dwelling, is subjected to the probability of finding another flat-mate with the same characteristics. If the individual is simulated to flat-mate with another individual, then the list of possible flat-mates is selected, which consists of all individuals of the same sex and age group who live on their own or flat-mate with others in a rented dwelling. A possible flat-mate is randomly selected and the size of their current households is examined.

If both individuals live with others, then a new household is created, which is denoted as a one with *medium search intensity*. The information concerning the flat-mates as well as their initial households is updated. If the initial households have only one member left, then the member's information is updated to denote that he/she no longer flat-mates (hld position = single). In addition, the initial households are denoted as households with *low search intensity.* It should be noted that the initial households will search for another dwelling only if the new household finds a new dwelling and consequently, the flat-mate formation actually takes place.

If only one of the flat-mates lives alone that is, he/she already has a dwelling of his/her own, then the other flat-mate moves in with him/her and becomes member of his/her household. The information concerning the flat-mates, the single's household as well as the other flat-mate's initial household is updated accordingly. In addition, both households are denoted as households with *low search intensity.*

If both flat-mates live alone that is, they both have their own dwellings, it is decided whose dwelling and consequently household the flat-mates keep. The two dwellings are compared based on size (number of rooms) and rent value. Once it is decided whose flat-mate dwelling is kept, the other flat-mate moves in with him/her and becomes member of his/her household. The information concerning the flat-mates as well as the household kept is updated accordingly. The household left has its ceased date set to denote that the household no longer belongs to the city's household set and the dwelling the household occupied is vacated and added to the housing stock. In addition, the household kept is denoted as a one with *low search intensity.*

5.3.7 Flat-mate Leaving event

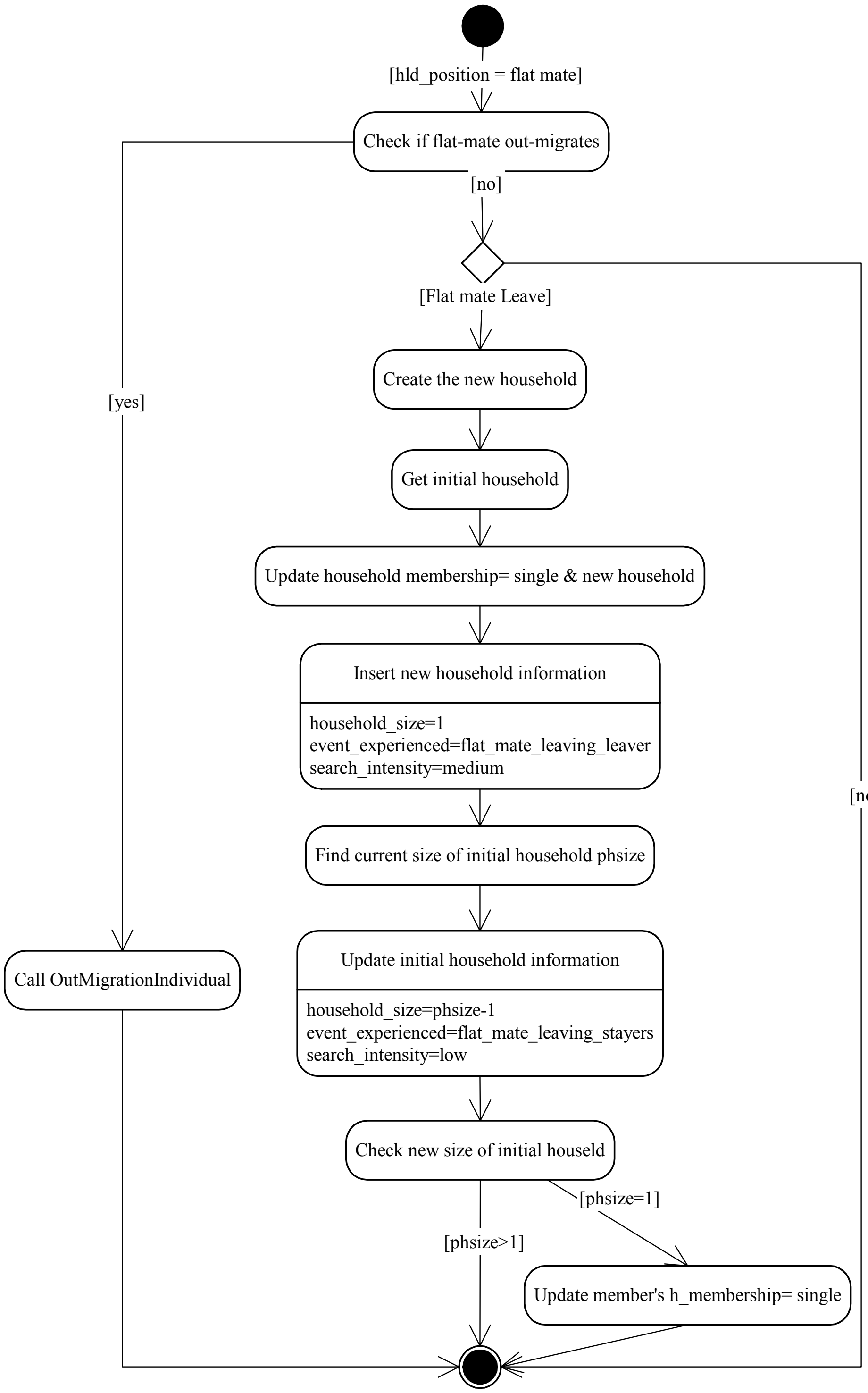


Figure 33: Flat-mate Leaving event

Each individual, who flat-mates with other individuals, is initially subjected to the probability of out-migrating. If the individual remains in the housing market system under examination then he/she is subjected to the probability of leaving the household to live on his/her own. If the individual is simulated to leave his/her household then he/she forms a new single-member household, which is denoted as a one with *medium search intensity*. The information concerning the individual as well as the initial household is updated. If the initial household has now only one member then his/her information is also updated. In addition, the initial household is denoted as one with *low search intensity.* It should be noted that the initial household will search for another dwelling only if the new household finds a new dwelling and consequently, the flat-mate leave actually takes place.

5.3.8 Out-migration event

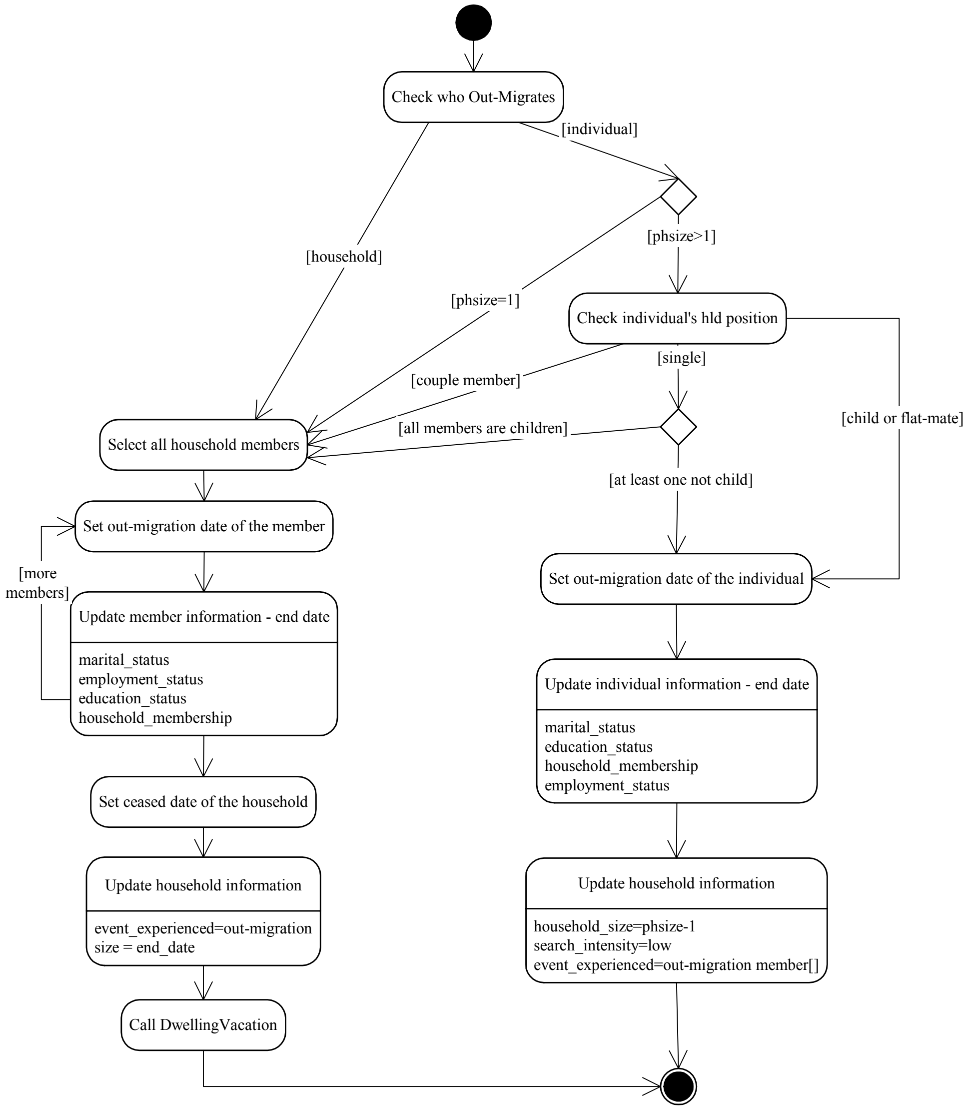


Figure 34: Out-migration event

A household or an individual is subjected to the probability of out-migrating. If a household is simulated to out-migrate then the out-migration date of each household member is set and his/her information is updated to denote that the individual no longer belongs to the population. In addition, the household's ceased date is set to denote that the household no longer belongs to the city's household set and the dwelling the household occupied is vacated and added to the housing stock.

If an individual is simulated to out-migrate, then the whole household to which the individual belongs out-migrates in the following cases:

- If the individual belongs to a one-member household.
- If the individual has a partner (he/she was a couple member).
- If the individual is the parent in a one-parent household with one or more children.

In all other cases the individual out-migrates on his/her own, that is, his/her out-migration date is set and his/her information is updated to denote that the individual no longer belongs to the population. In addition, the household's information is updated and it is denoted as a household with *low search intensity*.

5.3.9 In-migration event

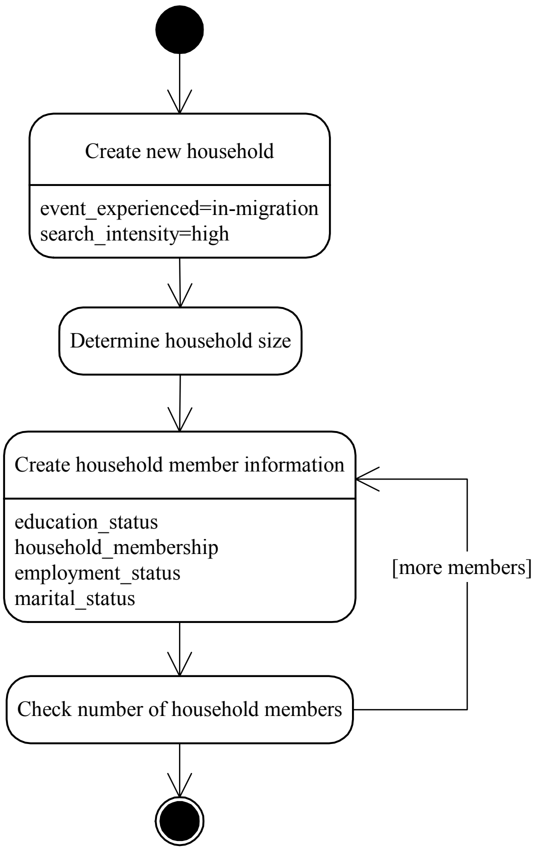


Figure 35: In-migration event

The number as well as the type of households entering the housing market is determined. For each in-migrating household a new household is created and denoted as a household with *high search intensity*. Next the number of members of the household is determined based on the household type. The individuals – members of the household are added to the population and information concerning each individual is inserted.

5.3.10 Education event

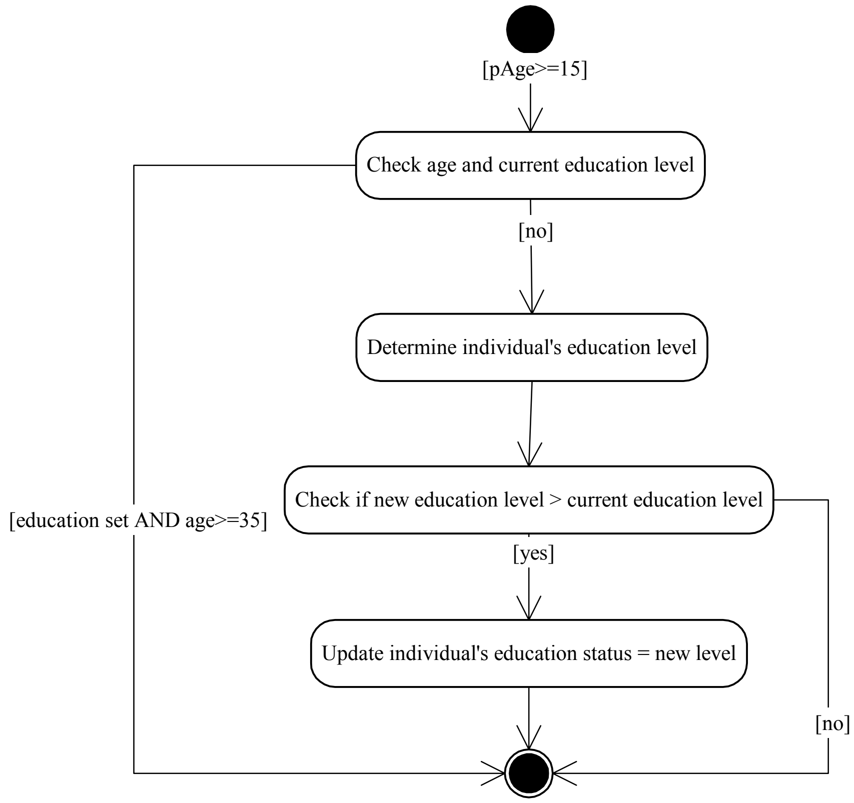


Figure 36: Education event

The education level of an individual may not change if he/she already has an education level and is over 35. In all other cases, each individual of age 15 and above is subjected to the probability of having his/her education level determined based on the individual's sex and age (none, primary or less, lower secondary, upper secondary, post secondary, tertiary). If the individual's new education level is higher than his/her current education level then the information concerning the individual is updated.

5.3.11 Job entry event

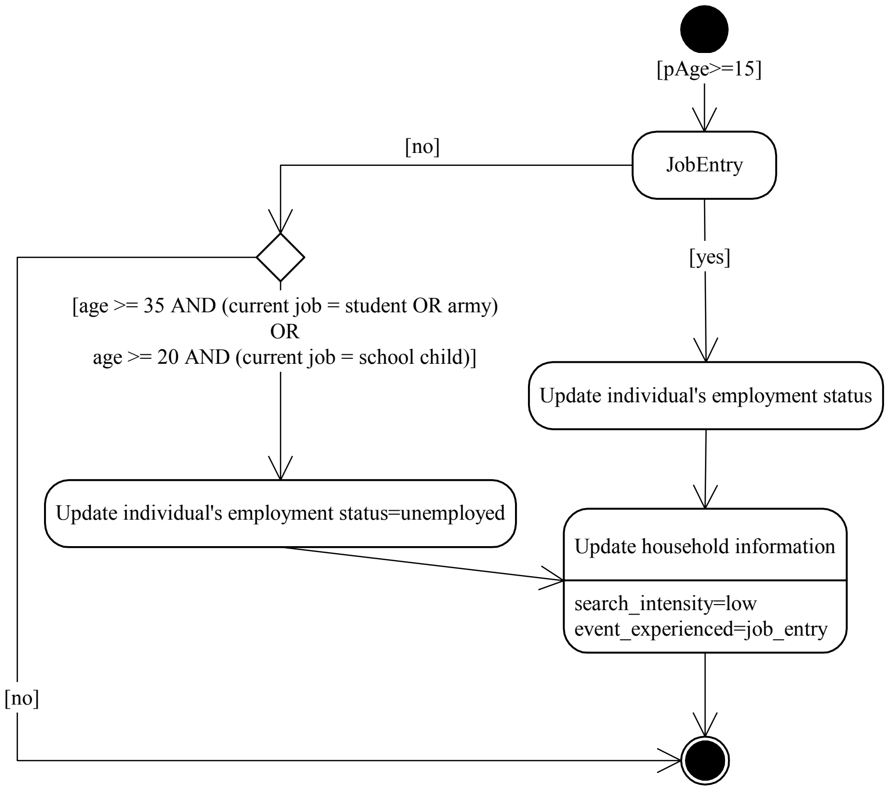


Figure 37: Job entry event

Each individual of age 15 and above, who does not work (soldiers are excluded) is subjected to the probability of getting a job based on the individual's sex, age, current employment (unemployed, student, school-child, army service, other inactive) and education status. If the individual is simulated to get a job (one of employee, employer, employed, family-worker, family-duties, and student) then the information concerning the individual as well as its household is updated and the household is denoted as one with *low search intensity*. If the individual is not simulated to get a job then if he/she is over 35 and a student or engaged in army service, or he/she is over 20 and a school-child then his/her employment status is changed to unemployed.

5.3.12 Job change event

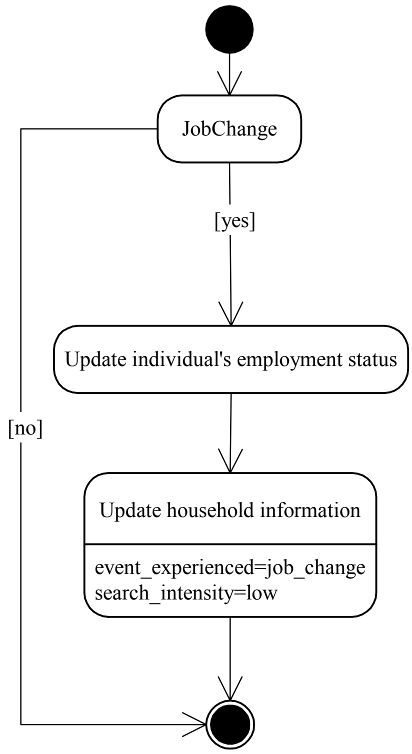


Figure 38: Job change event

Each individual who already works is subjected to the probability of job changing based on the individual's sex, age, current employment (employee, employer, self-employed, family-worker) and education status. If the individual is simulated to change job (one of employee, employer, employed, family-worker, unemployed, family duties, retired, other inactive) then the information concerning the individual as well as its household is updated and the household is denoted as one with *low search intensity*.

5.3.13 Demand change event

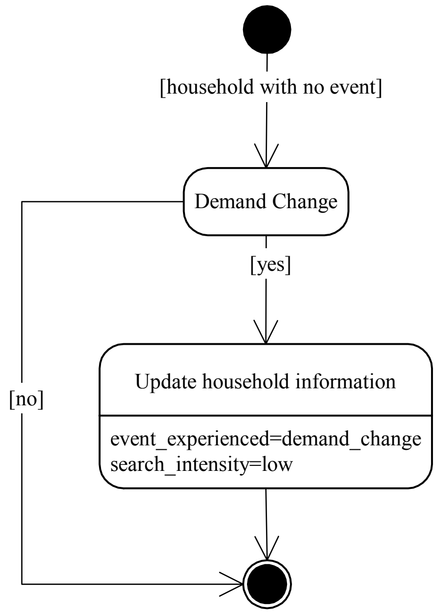


Figure 39: Demand change event

Each household that does not have experienced a demographic or income change event is subjected to the probability of searching for a new dwelling due to demand changing. If the household is simulated to have its demands changed then its information is updated and it is denoted as a household with *low search intensity*.

5.4 The Residential Search and Migration events

In this section each event of the *Residential Search and Migration* sub-model shown in Figure 26 is further elaborated.

5.4.1 Dwelling Search event

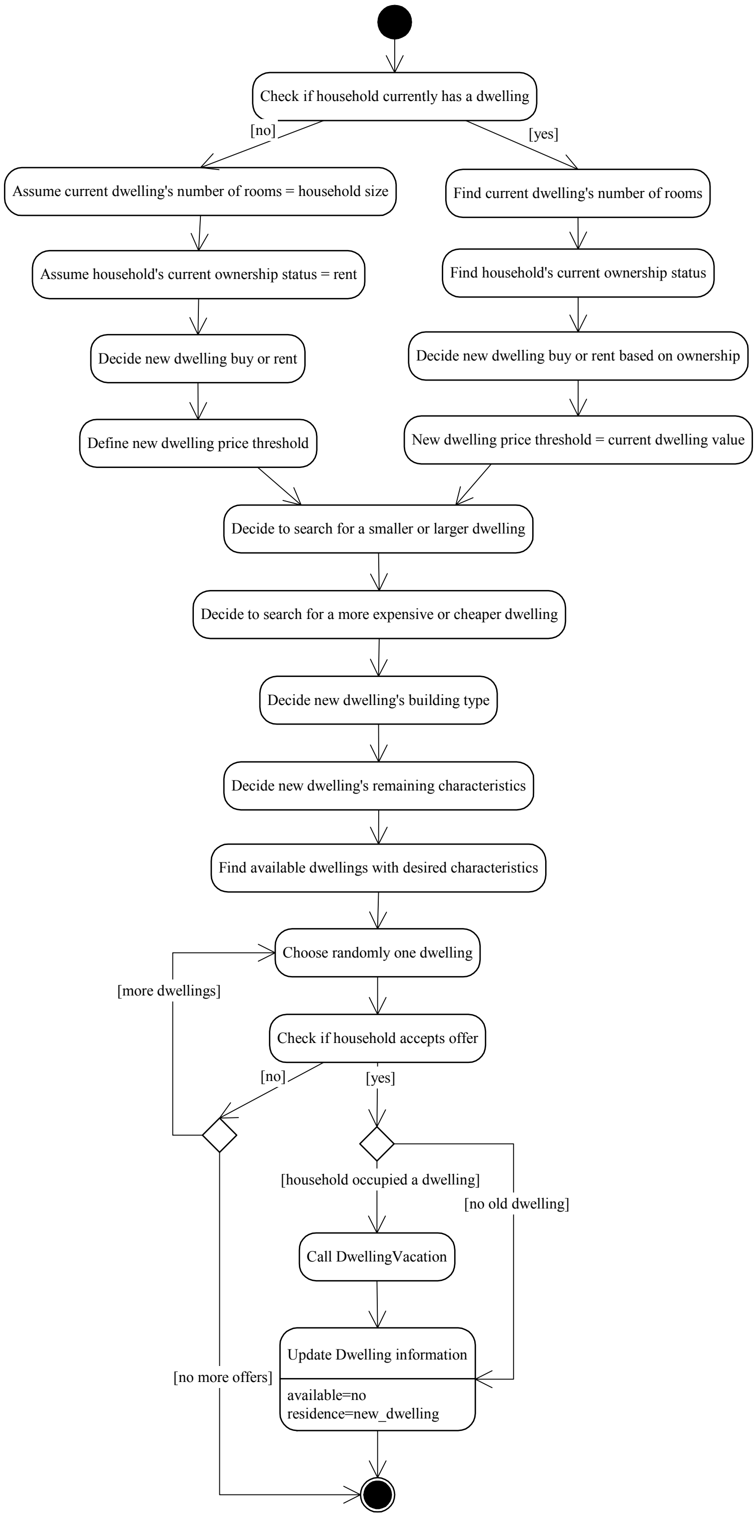


Figure 40: Dwelling search event

Each newly formed household or each household whose structure or housing circumstances have changed is simulated to search the list of vacant dwellings. Initially, it is decided whether the household will search for a dwelling to buy or rent based on the household's current ownership status, (the ownership status of a newly formed household is assumed to be renter). Next, the rent/market value threshold as well as the number of rooms threshold is defined based on the current dwelling's rent/market value and number of rooms (the rent/market value threshold of a newly formed household is given exogenously and the number of rooms is assumed to be equal to the household size). Based on these threshold values and the event experienced by the household, it is decided whether the household will search for a larger or a smaller as well as for a more expensive or a cheaper dwelling. In addition, the new dwelling's building type is determined. Next, the vacant dwellings with the desired characteristics are selected. If there is at least one available dwelling that matches the criteria, one dwelling is randomly chosen. If the household accepts the offer then the dwelling the household occupied (if any) is vacated and added to the housing stock and the new dwelling's information is updated. If the household rejects the dwelling then another suitable dwelling is randomly chosen and offered to the household successively until the household either accepts one dwelling or the number of dwellings the household may be offered is exceeded.

5.4.2 Dwelling vacation event

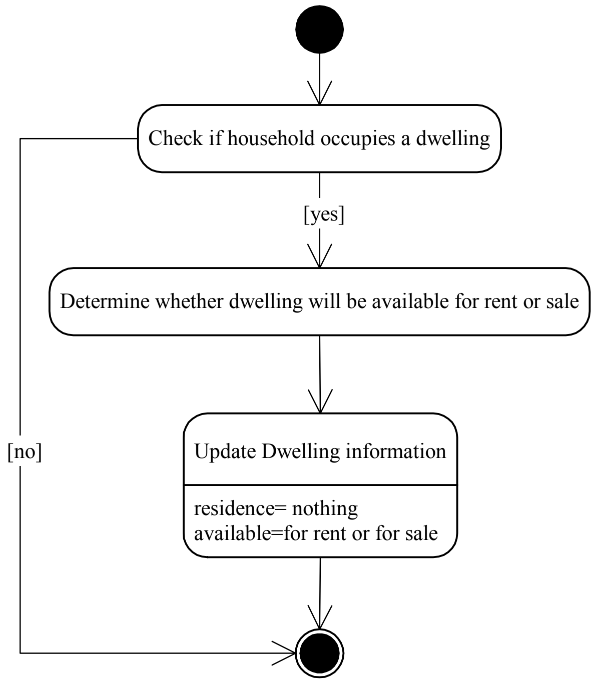


Figure 41: Dwelling vacation event

If the household moves to a new dwelling then if the household occupied a dwelling it should be vacated. Initially it is decided whether the dwelling will be available for rent or sale. Next, the dwelling's information is updated to denote that the dwelling is vacated and added to the housing stock.

5.5 The URM-Microsim Database

While the analysis model does not contain any implementation details in order to be usable by all members of a development team (users, analysts, programmers etc), the design model of a system should take into consideration the implementation platform to be used. Therefore, the database entities defined in the analysis model (in section 4.3), are further elaborated taking into consideration the implementation platform chosen. Since the ArcGIS platform is used for the implementation of *URM-Microsim*, the ArcInfo UML Geodatabase Model (an extension of UML) is used for the graphical representation of the design model.

In the *URM-Microsim* analysis model each database entity (*individual, household, dwelling, occupation, building, building block* and *residence zone)* corresponds to one class in the UML class diagram, which in turn, corresponds to a table in the ArcGIS database schema. Since ArcGIS platform does not allow the explicit modeling of the temporal semantics of information, each entity characteristic with temporal semantics in the analysis model is converted to a separate class in the design model. This class is related to the entity class via a composite relationship to denote that the temporal characteristic belongs to the entity.

For example, the *marital status* of an individual may change over time since an individual who is married may get a divorce, and may then get married again. Therefore, the *marital status* characteristic is modeled as a separate class (*marital status class*) with the following attributes:

- *marital status* which contains the marital status value of the individual
- *start_date* and *end_date* to denote the time period the individual has a single marital status value and

- *IndividualID* to denote the ID of the individual to whom the marital status value refers.

In addition, the marital status class is related to the Individual class via the Has_marital_status composite relationship.

In the design model the type of each entity characteristic (class attribute) should be defined based on the types available by the ArcGIS platform. For example the *birth_date* of an individual is of *esriFieldTypeDate* type. Moreover, the ArcInfo UML Geodatabase Model allows the use of *coded value domains* to give a more understandable description to the set of permissible values an entity characteristic may have. For example the *marital status* of an individual may be one of *never married, married, divorced, widow* or *cohabitating*. The type of marital status is defined in a coded value domain class called *marital status type*, according to which marital status is of *esriFieldTypeInteger* type and may have one of the 1,2,3,4 or 5 values (1 corresponds to never married, 2 corresponds to married, etc).

In the *URM-Microsim* design model seven (7) new class attributes were defined. These attributes are not entity characteristics and that is why they were not defined in the analysis model; they are in fact information that is required in the implementation of the design model. In other words, these attributes are programmatic attributes and are required in program execution.

- In *Individual* class the *event_in_process* attribute is defined to denote whether an individual is currently experiencing an event or not.
- In *Household size* class the *leaverID* attribute is defined to store the ID of the individual that left the household.

- In *Household* class the *num_of_cycles* attribute is defined to denote the searching cycle in which the household is. In *URM-Microsim* each household may search the housing market in four cycles.
- In *Event experienced* class the *hld_move* attribute is defined to denote whether the household has actually moved from one dwelling to another.
- In *Building block* class the *code_esye* attribute is defined to store the building block code given by the National Statistical Service of Greece (NSS).
- In *Building block* class the *Hld_change* attribute is defined to store the number of households that moved in – the number of households that moved out of the building block.
- In *Building* class the *code_esye* attribute is defined to store the building block code given by the National Statistical Service of Greece (NSS).
- In *Residence zone* class the *Hld_change* attribute is defined to store the number of households that moved in – the number of household that moved out of the residence zone.

Finally, it should be noted that each entity with non-spatial semantics such as individual is denoted as a subtype of the *Object* ESRI Class whereas each spatial entity such as building block is denoted as a subtype of the *Feature* ESRI Class.

Figure 42 shows the *Class diagram* corresponding to the design model of the *URM-Microsim* and Figure 43 shows the *classes* corresponding to the *coded value domains* defined in the design model.

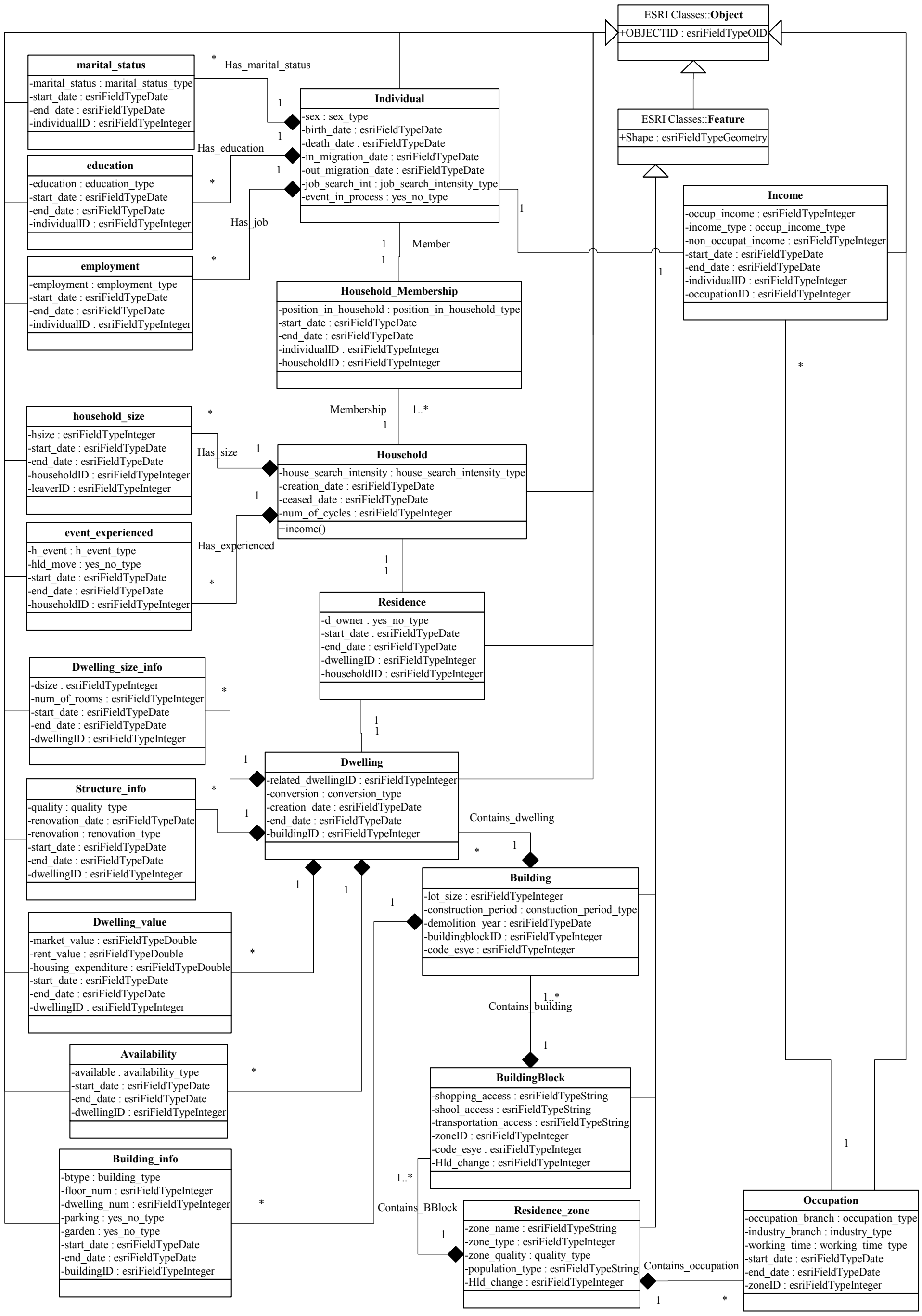


Figure 42: The URM-Microsim Database

Figure 43: The URM-Microsim Database Domains

5.6 Implementation evolution

Once the design model of *URM-Microsim* is completed, the implementation phase of system development is initiated. Since all diagrams are created using the Microsoft Visio 2000 UML CASE tool and will be implemented in ArcGIS, the class diagram corresponding to the *URM-Microsim Database* is initially exported to the Microsoft Repository by Microsoft Visio 2000. Then, the Schema Wizard in ArcGIS ArcCatalog is used to create the database schema. The database schema is stored in the *repository.mdb* Access file.

Once the database is created, then for each activity diagram presented in the previous sections, code is written in ArcInfo VBA and the program is called UrbanModel. In particular four modules are created:

- *HousingMarket*: it is the main module of the program and contains the main routine; from this routine the HousingDemand and SearchAndMigration modules are called (executed).
- *HousingDemand*: it corresponds to the Housing Demand sub-model of the *URM-Microsim* activity diagram (Figure 26). In addition, it contains all routines that simulate the demographic and income and employement events and correspond to the activity diagrams shown in Figure 27 to Figure 38**.**
- *SearchAndMigration*: it corresponds to the Residential Search and Migration sub-model of the *URM-Microsim Model* activity diagram (Figure 26). In addition, it contains the dwelling search routines that correspond to the activity diagrams shown in Figure 40 and Figure 41.
- *UrbanModelLibrary*: it contains subsidiary routines such as routines for accessing the temporal information of entities (MaritalStatusInsert, MaritalStatusUpdate), a routine for calculating the minimum value of ten values (Min_Value) etc.

It should be noted that during the course of model development, parameter estimation and model application, a number of *URM-Microsim* elements were under- or undeveloped. First, the Housing Supply sub-model was not implemented. All new dwellings and structure conversions are given exogenously. In addition, rent and market value changes of the dwellings are also given exogenously. Second, the income event was not implemented and households do not search for a new dwelling due to income changes. Third, in job entry and job change events the location of employment is not taken into consideration. Finally, in-migration event takes into consideration only one type of in-migrating households: students.

The program may be executed by clicking the *Housing_Market* button added at the tools toolbar.

5.7 Summary

In this chapter, the design model of *URM-Microsim* was presented taking into consideration the implementation platform to be used. Therefore, UML with ArcInfo extensions (ArcInfo Geodatabase Model) in conjunction with the Microsoft Visio 2000 UML CASE tool is used for the graphical representation of the design model. Initially, the behavior of the *URM-Microsim Model* and its sub-models *(Housing Demand, Housing Supply* and *Residential Search and Migration)* was presented, that is, the steps the model goes through, in order to execute the model processes. Next, the events that occur in the Housing Demand and the Residential Search and Migration sub-models were presented. In addition, the *URM-Microsim* Database was presented, namely, the *entities* that are of interest as well as *their characteristics* and the *relationships* with each other. Finally, implementation issues concerning the implementation of the *URM-Microsim* design model were discussed.

Chapter 6
Data and Synthetic Population Development

6.1 Introduction

This chapter is devoted to applying *URM-Microsim* with real world data. The Municipality of Mytilene was chosen as a testing ground for *URM-Microsim*. Mytilene was chosen for two practical reasons. First, it is a small town (about 30.000 population) with a single economy and quite self-contained. Secondly, since it is the seat of the Department of Geography of the University of the Aegean, it was reasoned that it would be easier to collect the required data and in cases where no data are available, realistic assumption could be made. Section 6.2 presents a short introduction on the history and background of Mytelene, its population and its housing market. Section 6.3 presents the population synthesis technique used to produce the Mytilene data as well as its input data and gives an excerpt of the synthetic data. Finally, in section 6.4 the probability distributions required by *URM-Microsim* are calculated.

6.2 The Mytilene Housing Market

Mytilene is built on the southeastern coast of Lesvos island (Figure 44 and Figure 45), one of the islands of the Aegean Sea. It is the largest city and the capital of the Prefecture of Lesvos, as well as the administrative and economic centre of the island. The urban planning and the architecture of the city reflect the needs for spatial organisation according to the political, economical and social conditions that prevailed in the past.

Figure 44: Map of Lesvos Island

Figure 45: The location of the Municipality of Mytilene

6.2.1 History

From the mid 19th century until the interwar period (1920-1940), Mytilene experienced important economic and cultural development (Βουνάτσου, 2002). The economy of Mytilene was based on olive oil production and trade. Trade activities were extended in all corners of the Mediterranean Sea and the Black Sea. Historical events such as the *Balkan Wars, World War I,* the *Russian Revolution* and the *Near East Catastrophe* as well as the shifts of the commercial axes altered the favourable economic conditions and set the stage for a prolonged recession. After 1912, the city experienced demographic growth due to the liberation of Mytilene from the Ottoman Empire and the exchange of populations that followed the Near East Catastrophe in 1922. At that time the city of Mytilene embraced 10.000 refugees causing a significant increase in housing demand, which in turn set the stage for the reconstruction of the city. At the same time, the transactions and the communication with Near East were interrupted, with unfavourable repercussions in the economy of the whole island (mainly in the trade) and the demographic characteristics. In 1928, the city of Mytilene had 27.870 inhabitants, in 1940 it had 24.351 inhabitants, and in 1951 it had 28.837 inhabitants.

Thus, from 1912 to 1960 a progressive reconstruction of the city is observed and the first high rise buildings were erected (Βουνάτσου, 2002). The city was reorganized and its reconstruction and growth was initiated. Furthermore, after 1923 the image of the city changed due to the withdrawal of the Muslims and the settlement of the refugees. New districts were built and titles of property were conveyed. Despite the repercussions of the Near East Catastrophe in the economic and social life, commercial activity and industrial production flourished up until 1940.

On the contrary, from 1960 to 1974, significant drop in olive oil prices created and economic crisis accompanied by the massive out-migration of population to foreign countries (Germany, U.S.A., Australia and Canad) and large urban centres (Athens and Thessalonica) (Βουνάτσου, 2002). In 1974 the crisis accelerated, owing to the Greek-Turkish crisis and the reconstruction of the city was held back. After 1974, the reconstruction of Mytilene restarted and households left old houses to live in new apartments in blocks of flats. However, the city's growth was hampered during the 1980s because of the prevailing high interest rates, and was significantly increased during the 1990s.

In particular, during the 1990s significant public works were financed by the European Union. Such is the construction of the city's biological infrastructure and Marina, and the introduction of new technologies in critical sectors of the public administration, aiming to provide improved services to the citizens. Moreover, the governmental policy of decentralisation led to the establishment of the University of the Aegean with six of its Departments and the University's headquarters settled in Mytilene. In addition, the governmental policy of decentralised public administration led to the enactment of the North Aegean Prefecture as well as the Ministry of the Aegean with their headquarters settled in Mytilene. These governmental policies activated the development of the city since they created new places of work and restrained population out-migration, evoked the in-migration of students and therefore contributed in the economic growth and prosperity of the city. New building construction as well as old building conversion was significantly increased to meet the increase in housing demand (Βουνάτσου, 2002).

6.2.2 Residence zones

The continuous demographic and economic growth of Mytilene during the 20th century, led to the extension of the residential space to the rural regions of city, as well as to the configuration of industrial zones round the harbours. The extension was focused along the north-south axis, the coast and the two harbours, but the residential development also included the hills at the western boundaries of the city. Commercial, craft-based, social and symbolic centre of the city constitutes the Market Street (Ermou Street) that is extended from north to south connecting the two harbours. The economic growth of the city kept up with the growth of the southern harbour of the city, while the northern harbour that was used from the ancient years, was abandoned in the 18th century, due to the illuviations that did not allow the mooring of boats.

The residential growth of Mytilene led to the creation of new districts and to the continuous change of the urban network. However, most districts maintain up to date the same boundaries and name since the beginning of the 20th century. The most important residence zones of Mytilene *(Area 1)* starting from north are the following:

- The *Sinikoismos* residence zone (zoneID= 13) is the district where the refugees who arrived in Lesvos Island in 1922 settled. The residence zone was established in 1929 in order to cope with the massive needs of the refugees for accommodation and takes up the hills westwards of the Northern harbour.
- The *Paidikos Stathmos* residence zone (zoneID=12) lies next to the Sinikoismos district and received many refugees in 1922.
- The *Epano Scala* residence zone (zoneID= 14) is located at the Northern harbour and takes up the northern part of the Ermou market street, which constituted the old Turkish Market. It consisted of the houses of the

Muslim residents of the city, as well as their cemeteries and temples, and received many refugees in 1922.

- The *Agios Georgios* residence zone (zoneID= 15) extends round the homonym temple at the west side of the Ermou market street.
- The *Agias Eirinis* (zoneID= 9), *Kentriki Agora* (zoneID= 10) and *Kamares* (zoneID= 11) residence zones mainly consist of commercial shops, craft-based shops and offices in the Ermou and Kavetsou streets and include few residences.
- The *Koulbara* residence zone (zoneID= 21) is a large district that extends towards the west side of the Castle and east of the Ermou market street. It includes a large number of small neighbourhoods, with dense residential layout.
- The *Kioski* residence zone (zoneID =22) is located at the hills south of the Koulbara district and the Castle. It constituted one of the aristocratic districts of Mytilene and included many mansions that are maintained up to date.
- The *Lagada* residence zone (zoneID =8) that takes up the western entry of the city. The district is considerably under populated with small and low cost houses.
- The *Vounaraki* residence zone (zoneID =7) was developed at the beginning of the 20th century at the hill south of the Lagada district.
- The *Makris Yalos* residence zone (zoneID =6) was initially under populated and consisted of many mansions along the seacoast with large courtyards and gardens. After 1960, the residences' owners sold the courtyards and gardens to contractors who built large apartment blocks.
- The *Sourada* (zoneID =18) and *Akelidiou* (zoneID =19) residence zones used to be the resorts of the city's upper class and henceforth have been included in the urban network. They were relatively under populated regions, with impressive mansions with large gardens and impressive

views of the sea. After 1980, the residences' owners sold the gardens to contractors who built semi-detached double-deck houses and maisonettes.

- The *Chrisomalousa* (zoneID = 1), *Kalithea* (zoneID = 2), *Gipedo* (zoneID = 3), *Pirgelia* (zoneID = 4), *Maura Sidera* (zoneID = 5), *Nosokomeio* (zoneID = 16), *OEK Chrisomalousa* (zoneID = 20) residence zones are over-populated districts at the south-west part of the city. Up to the mid of the 20th century these districts were regions with gardens, while in the past decades they developed rapidly with the construction of large blocks of flats.
- The *Halikas*(zoneID =17) residence zone used to be a suburb of Mytilene, is currently included in the urban network of city, and extends in the slope over the Chrisomalousa district.
- The *Vareia* (zoneID =23), *Neapolis* (zoneID =24) and *Pligoni* (zoneID =25) residence zones are southern suburbs of Mytilene, that experience a rapid growth in the last decade since they have a large number of developable blocks and many maisonettes or semi-detached double-deck houses are being constructed.

Additionally, a few more residence zones at the southern suburbs of Mytilene experience a rapid growth during the last decade due to the developable blocks they have:

- The *Agia Marina* (zoneID =26) *Agia Paraskeui* (zoneID =27), *Kratigos* (zoneID =28) residence zones (*Area 2*).
- The *Taksiarxon* (zoneID =29) residence zone (*Area 9*).

Finally, is should be noted that the Municipality of Mytilene includes the Alyfanta (*Area 3*), Afalona (*Area 4*), Loutra *(Area 5*), Moria (*Area 6*), Pamfila (*Area 7*) and Panagiouda (*Area 8*) villages that are outside the scope of this work since they are not included in the urban web.

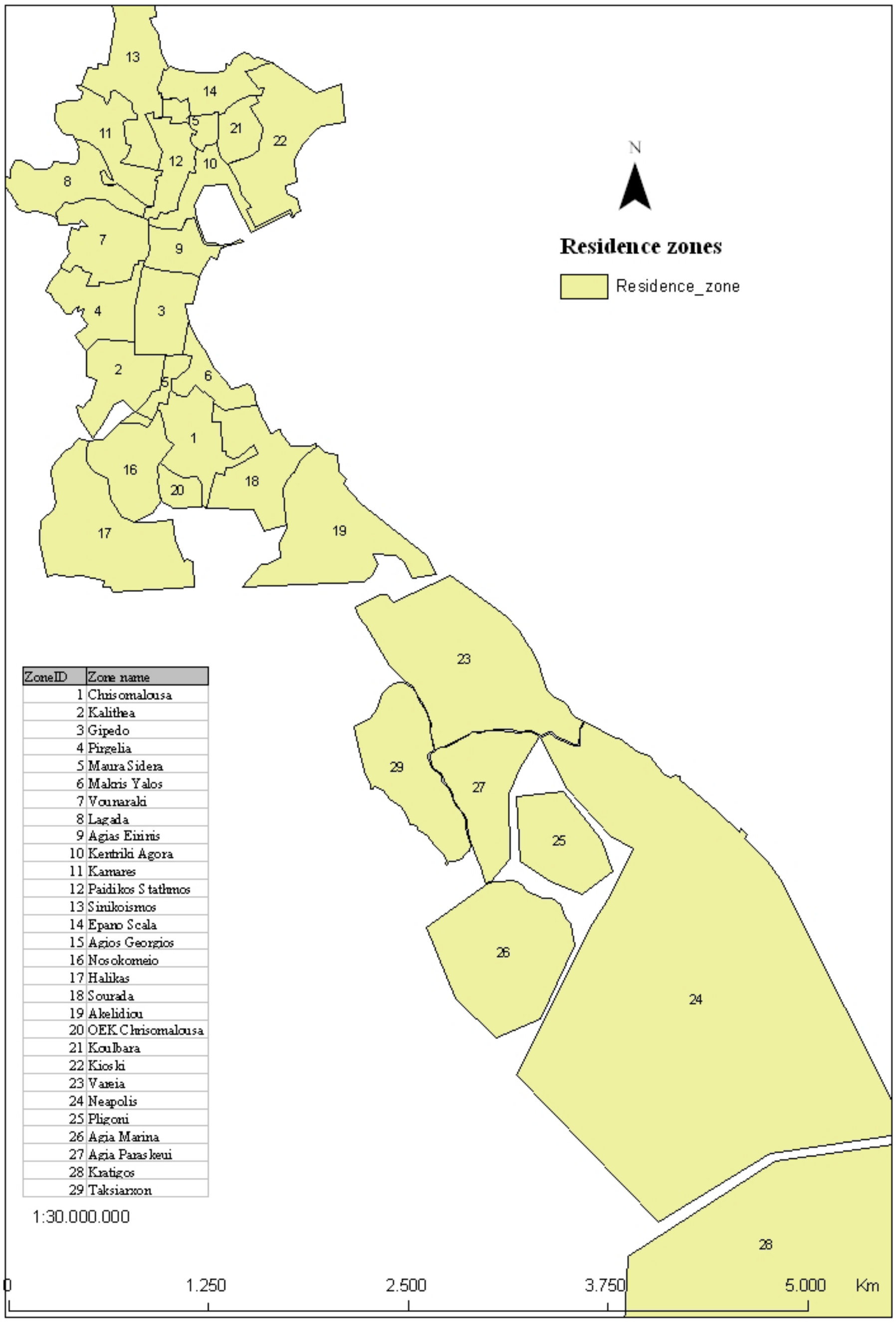


Figure 46: Mytilene residence zones

6.3 Population and Housing Stock Data

URM-Microsim is based on a dynamic, spatio-temporal microsimulation model that requires micro level input data. Since population micro data is suppressed to maintain confidentiality, a population synthesis technique is used as a viable alternative to the collection of micro data (Ryan *et al.*, 2006).

6.3.1 Population synthesis technique

Population synthesis techniques are algorithms that take aggregate population data, as well as sample population data as inputs, and produce a complete list of a population's members, each with associated attribute data, as output (Ryan *et al.*, 2006). Various types of synthetic populations can be created to suit different needs such as individuals, households and dwellings. The basic idea behind population synthesis is that given whatever population information is available, a list of members of the population is created, using some algorithm, such that the synthetic population conforms to the base information. Then it can be said that the synthetic population is among the set of 'best possible' estimates of the actual population, given the input information. Of course, different algorithms may produce synthetic populations, which all conform to the input data, but differ in their quality. This is the result of each algorithm having its own underlying theory, which may or may not be sound. A common situation is to have input data consisting of a *small sample* from the population (generally with no spatial identifiers), as well as *tabulations* representing the distribution of population characteristics over space. It is to this situation that the Combinatorial Optimisation (CO) technique (Voas and Williamson, 2000), which is used to produce the data of Mytilene, is particularly well suited.

In the Combinatorial Optimization (CO) method, the study area for which a population is to be created is normally divided into mutually exclusive and exhaustive zones such as census tracts, enumeration areas or traffic analysis zones. If each member of the synthetic population is to contain certain attribute variables, then *two forms of input data* are required by the CO method. First, *a micro sample* from the whole population is needed, where members of the sample contain the attribute variables desired in members of the synthetic population. Second, *tabulations for a sub-set* of the desired attribute variables are required, showing the distribution of those variables over the zones comprising the study area. For example, if a synthetic population is to be created with members having the characteristics of sex, age and height, then the sample members must have values for each of these variables, and tabulations for at least one of the variables is also required.

The CO method creates a population for each zone in the study area separately, by fitting a sub-set of the sample to the tabulations for each zone. First, a randomly selected sub-set of individuals from the sample is selected, matching the population size of the zone. A statistic is calculated to measure the fit of this sub-set to the known distributions of characteristics in the zone. Following this, one of the individuals from the sub-set is switched randomly with another individual from the sample (with replacement), and the statistic is calculated again. If the overall fit of the new sub-set is superior to that of the original sub-set, then the switch is maintained; otherwise, the original sub-set is maintained. This process is repeated until a threshold value of the comparison statistic is reached, or a user defined iteration limit is reached.

A program to execute the CO method was written in C++ and is called the Synthpop program.

6.3.2 Base Data

A variety of data sources were used to derive the inputs to the SynthPop program:

1. In order to collect the *micro sample population data*, a questionnaire was composed within the context of this work based on the questionnaire used during the National Statistical Census 2001. The questionnaire is presented in **Appendix D**. The data was collected by students who attended the Housing Geography undergraduate course of the Department of Geography of the University of the Aegean (April – May 2005). The study area was Chrisomalousa (residence zoneID =1). It should be noted that 109 questionnaires were filled in, that is, 109 households corresponding to 280 individuals. It is a very small sample since it corresponds to the 0,82% of the city's population.
2. The *tabulations* representing the distribution of population characteristics over space were derived from the National Statistical Service of Greece (NSS) data tables based on the National Statistical Population Census 2001 data for the Municipality of Mytilene. In addition, the *tabulations* representing the distribution of building characteristics over space were derived from the NSS data tables based on the National Statistical Census 2000 data for the Municipality of Mytilene (http://www.statistics.gr/). The *tabulations* that were used as inputs to the population synthesis technique are shown in Table 3 to Table 13.
3. The rent and market values of dwellings were derived from a real estate agency research regarding the city of Mytilene (Βουνάτσου, 2002) (Table 14).
4. The ArcInfo *feature table* representing the building block spatial characteristics was produced by the Spatial Analysis, GIS and Remote Sensing Laboratory of the Department of Geography of the University of the Aegean (Table 15).

5. The ArcInfo *feature table* representing the building spatial characteristics was derived from the National cadastre AUTOCAD layers (Table 16).

GEO CODE	PREFECT, AREA	CENCUS CODE	BLOCK CODE	USUAL RESIDENCE POPULATION	Private Households	Members	Dwellings
83010000	**MUNICIPALITY OF MYTILENE**			**37.881**	**13.227**	**34.465**	
83010100	**Area of Mytilene (Area 1)**			**30.576**	**10.877**	**28.168**	
83010101	**Mytilene (Sub-Area 1.1)**			**28.950**	**10.330**	**26.676**	**15.146**
83010101		028	0000	0	0	0	6
83010101		028	0002	4	3	4	4
83010101		028	0005	0	0	0	1
83010101		028	0007	5	2	5	5
83010101		028	0008	33	11	33	17
83010101		028	0010	23	9	22	16
83010101		028	0011	9	5	9	8
83010101		028	0012	18	5	17	12
83010101		028	0013	31	9	30	12
83010101		028	0014	47	19	46	20
83010101		028	0015	12	6	12	11
83010101		028	0016	60	23	58	31
83010101		028	0017	16	7	17	13
83010101		028	0018	43	13	41	23
83010101		028	0019	37	15	37	24
83010101		028	0020	44	12	36	13

Table 3: Excerpt from households, household members and dwellings per building block

AgeCat	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Grand Total
0-14	4993	71	44	68	198	205	191	114	42	5926
15-24	4647	128	303	50	248	492	148	87	55	6158
25-39	6973	137	133	97	274	287	251	151	61	8364
40-54	6018	134	68	69	251	245	269	134	70	7258
55-64	2873	100	41	66	148	133	183	75	36	3655
65-79	3967	106	42	91	270	237	230	115	72	5130
80+	1105	24	11	24	61	75	57	22	11	1390
Grand Total	30576	700	642	465	1450	1674	1329	698	347	37881

Table 4: Individuals by age categories and area

Sex	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Grand Total
male	14818	395	483	241	758	1025	687	341	177	18925
female	15758	305	159	224	692	649	642	357	170	18956
Grand Total	30576	700	642	465	1450	1674	1329	698	347	37881

Table 5: Individuals by sex and area

Industry Sectors	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Grand Total
Primary (NACE groups A/B)	498	42	14	36	191	51	90	23	13	958
Secondary (NACE groups C/F)	1672	46	23	48	125	75	68	39	35	2131
Tertiary (NACE groups G/Q)	7856	139	101	51	195	181	231	132	45	8931
Not Specified	862	31	8	10	12	40	50	10	10	1033
Unemployed	1260	28	17	5	47	51	50	29	19	1506
Grand Total	12148	286	163	150	570	398	489	233	122	14559

Table 6: Employed individuals by occupation per area

Number of Members	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Grand Total
1	2609	37	20	40	104	109	117	46	25	3107
2	3093	71	39	62	159	142	134	73	32	3805
3	2453	32	25	36	100	101	108	57	29	2941
4	1968	47	19	33	95	85	93	54	23	2417
5	515	15	5	3	35	29	23	13	10	648
6	175	5	4	4	13	7	10	3	4	225
7	31	1	0	2	3	4	1	1	0	43
8	17	0	0	0	1	2	3	0	0	23
9	7	0	0	0	2	0	0	0	0	9
10	9	0	0	0	0	0	0	0	0	9
Grand Total	10877	208	112	180	512	479	489	247	123	13227

Table 7: Households by number of household members and area

Rooms	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Grand Total
1	375	0	19	3	78	18	19	12	4	528
2	1341	27	28	11	119	67	70	37	41	1740
3	3648	86	68	58	172	127	133	66	59	4416
4	5816	134	55	98	276	202	246	120	54	7000
5	5038	130	65	103	172	327	360	136	27	6358
Grand Total	16218	377	235	273	817	741	828	371	185	20045

Table 8: Dwellings by number of rooms and area

Use	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Grand Total
Occupied	10324	205	109	180	502	474	485	240	122	12641
Temporarily vacant	2172	8	17	15	35	32	68	24	3	2374
Vacant Seasonal	1548	151	84	46	183	180	150	79	41	2462
Vacant 4 rent/sale	2134	13	21	32	96	55	125	20	18	2514
non conventional dwellings	3	0	0	0	0	0	0	0	0	3
collective living quarters	37	0	4	0	1	0	0	8	1	51
Grand Total	16218	377	235	273	817	741	828	371	185	20045

Table 9: Dwelling use by type of use and are

Floors	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Grand Total
1	3302	315	205	213	653	442	672	235	80	6117
2	5577	217	110	193	175	429	442	148	115	7406
3	955	4	5	21	141	21	31	1	10	1189
4 to 6	345	0	1	1	2	1	2	0	1	353
7plus	14	0	0	0	0	0	0	0	0	14
Grand Total	10193	536	321	428	971	893	1147	384	206	15079

Table 10: Buildings by number of floors and area

Ground Floor Parking	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Grand Total
No	10106	535	320	428	971	893	1124	383	203	14963
Yes	87	1	1	0	0	0	23	1	3	116
Grand Total	10193	536	321	428	971	893	1147	384	206	15079

Table 11: Buildings by parking availability and area

Type	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Grand Total
RtoD(1)	5933	366	159	260	636	620	724	247	161	9106
RtoD(2)	1242	22	13	8	72	33	42	52	13	1497
RtoD(3to5)	658	1	2	11	5	8	5	8	3	701
RtoD(6to15)	212	0	0	3	0	0	1	0	0	216
RtoD(16to30)	45	0	0	0	0	0	0	0	0	45
RtoD(31plus)	2	0	0	0	0	0	0	0	0	2
RtoChurches	67	12	12	12	26	17	13	8	1	168
RtoHotels	20	1	19	0	3	0	1	2	0	46
RtoFactories	47	0	7	2	9	25	20	0	0	110
RtoSchools	39	1	2	1	2	2	5	4	1	57
RtoShops/Office	751	6	30	0	23	35	42	13	0	900
RtoCar Parks	0	0	0	0	0	0	0	0	0	0
RtoHospitals	1	0	0	0	1	1	0	0	0	3
RtoOther	504	119	66	125	141	99	244	32	24	1354
MtoD(1)	216	3	4	1	31	21	24	11	1	312

Type	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Grand Total
MtoD(2)	82	0	0	0	5	2	4	1	0	94
MtoD(3to 5)	65	0	0	0	0	0	0	3	0	68
MtoD(6to 15)	57	0	0	0	0	0	0	0	0	57
MtoD(16to30)	26	0	0	0	0	0	0	0	0	26
MtoD(31 plus)	10	0	0	0	0	0	0	0	0	10
MtoChurches	8	2	0	0	0	0	3	0	0	13
MtoHotels	39	0	0	0	5	1	0	0	0	45
MtoFactories	5	0	2	0	0	4	0	0	0	11
MtoSchools	3	0	0	0	0	1	0	0	0	4
MtoShops/Office	111	0	2	2	5	11	5	1	0	137
MtoCar Parks	0	0	0	0	0	0	0	0	0	0
MtoHospitals	1	0	0	0	0	0	0	0	0	1
MtoOther	49	3	3	3	7	13	14	2	2	96
Grand Total	10193	536	321	428	971	893	1147	384	206	15079

Table 12: Buildings by building type and area

Construction Period	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Grand Total
Before 1919	1223	45	9	60	55	241	438	8	2	2081
1919to1945	3126	55	33	171	307	298	245	99	30	4364
1946to1960	1386	40	31	36	129	113	126	57	52	1970
1961to1970	1118	58	23	43	165	62	127	37	36	1669
1971to1980	1088	112	58	28	124	33	50	44	22	1559
1981to1985	662	78	38	18	36	42	26	25	13	938
1986to1990	467	40	26	18	39	38	47	36	14	725
1991to1995	436	32	53	24	40	21	46	31	16	699
1996plus	459	46	25	15	27	27	18	33	12	662
Under construction	171	27	25	15	2	13	23	13	9	298
Not specified	57	3	0	0	47	5	1	1	0	114
Grand Total	10193	536	321	428	971	893	1147	384	206	15079

Table 13: Buildings by construction period and area

Number of rooms & type specific Rent & Market values per area						
Residence zones	**Number of rooms**					**Market value**
	1	**2**	**3**	**4**	**5 or more**	
7, 8, 10 - 15, 21, 22 (cheap areas)	260 euros	280 euros	300 euros	350 euros	370 euros	910 euros/m2
1 - 5, 16, 20, 25	280 euros	300 euros	350 euros	370 euros	400 euros	1200 euros/m2
0, 9, 17 - 19, 23, 24, 26 – 29 (expensive areas)	300 euros	350 euros	400 euros	420 euros	440 euros	1450 euros/m2
One dwelling house				450 euros		1500 'euros/m2

Table 14: Number of rooms & type specific Rent & Market values per area

OBJECTID	code_esye	zoneID
1	425	2
2	424	2
3	423	2
4	426	2
5	451	2
6	428	2
7	452	2
8	449	2
9	473	2
10	463	6
11	509	16
12	599	1
13	496	1
14	483	1
15	605	1
16	505	1
17	504	1
18	519	1
19	518	1
20	497	16

Table 15: Excerpt from Building blocks ArcInfo feature table

OBJECTID	Num_of_Floors	BLOCK_CODE
1	00	212
2	00	212
3	00	212
4	00	212
5	00	212
6	00	212
7	00	212
8	01	212
9	00	25
10	00	60
11	00	25
12	00	25
13	00	25
14	01	25
15	00	25
16	00	25
17	00	25
18	00	25
19	00	25
20	01	25

Table 16: Excerpt from Buildings ArcInfo feature table

6.3.3 Synthetic Data

The output of the SynthPop program is five tables: *Individual* (Table 17), *Household* (Table 18), *Dwelling* (Table 19), *Building* (Table 20) and *Building block* (Table 21). The population data created refer to the areas of Mytilene (Area 1), Agia Marina (Area 2) and Taxiarxon (Area 9) and includes 29085 individuals. Population is composed of individuals linked to households linked to dwellings linked to buildings linked to blocks.

Initially, the Combinatorial Optimization (CO) method was used to create the population of *individuals* from the micro sample population data (size = 280). The tabulations which were used to constrain the synthesis process were: age X area (Table 4), sex X area (Table 5) and household heads X area (Table 7). Once the population of individuals was created, it adopted additional attributes from the sample. The attributes of individuals are shown in Table 17.

Then the CO method was used to create the population of *households* from the micro sample population data (size = 109). The tabulation used to constrain the synthesis process was: number of household members X area (Table 7**)**. Once the population of households was created, it adopted additional attributes from the sample. The attributes of households are shown in Table 18.

Next, the list of *dwellings* was created based on the list of the number of dwellings, individuals, households and buildings per block (Table 3), the main use of each building (Table 12) and the number of floors in each building (Table 16). Each dwelling was assigned semi-randomly to a building on its block, where buildings having more floors had a correspondingly higher probability of being assigned dwellings. Based on the number of dwellings in a dwellings building, the number of rooms was assigned to each dwelling. The rules for assigning rooms to dwellings were as follows:

a. A building with one dwelling may have a dwelling with 4 to 7 rooms.
b. A building with two dwellings may have dwellings with 3 to 5 rooms.
c. A building with three or more dwellings may have dwellings with 1 to 4 rooms.
d. Based on the number of rooms and the area, each dwelling was assigned a rent and market value. (Table 14).

Buildings were assigned dwellings as described above. The Type attribute of building denotes the number of dwellings in case of a residential building and its general use in case of a non-residential building (for example Shops/Office) (Table 20).

Block data was entirely provided by the National Statistical Service of Greece (NSS).

Each individual from the synthesized population of individuals was linked to a household from the synthesized population of households. This was done in accordance with a set of rules to create "reasonable" households. The method begins with one-member households and then sequentially moves through the rest of them as follows:

1. If the Household of an area has only one member then an Individual from the area is chosen such that he/she is a household head and is never married, widow or divorced, and he/she is assigned to the Household.
2. If the Household of an area has two members then an Individual from the area is chosen such that he/she is a household head (his/her marital status is of no importance). Next, the second Individual is chosen such that if the household head is married or cohabitating then the second member has the same marital status with the household head, is of opposite sex, has an age category within (head age category-1, head age category+1) and has a husband/wife or cohabitee relation to the household head accordingly. If household head is not married or cohabitating then the next rule applies.
3. Rules for assigning 3rd, 4th etc. members to households where the Household Head is married or cohabitating OR for assigning 2nd, 3rd etc. members to households where Household head is not married or cohabitating. The new member:

a. Can not have a husband/wife or cohabitee relation to the household head
b. If the new member has a child relation to the household head then he/she must be one or two age categories younger than the youngest parent. In addition, the child can be married only if this individual will not be the last one selected for the family, i.e. other individuals are remained to be selected after this one.
c. If the new member has a grandson relation to the household head, then he/she must be two or more age categories younger than the youngest grandparent.
d. If the new member has a son in law relation to the household head, then there must already be a married daughter (ie. Female Child) in the Household. (Vice versa for Daughter in law).

The general rule is to assign as many of the original list of individuals to households as possible (without replacement), and then if the process gets stuck near the end of the assignments, another member(s) is taken from the original list (with replacement), assigning it a new individual ID.

Each household from the synthesized population of households was randomly linked to a dwelling from the population of dwellings.

Once the population data were created, a new module was written in VBA and added in the UrbanModel ArcInfo program called DataInsert. The module includes all routines necessary to load the data contained in the aforementioned data tables in the database created based on *URM-Microsim.* These routines highly depend on the data model used as well as the input data.

Indiv ID	Hld ID	Sex	Age Cat	Hld Head	Relation to Hld Head	Marital Status	Education Level	Employment Status	Occupation
1	1	'male'	'3'	'NO'	'Husband	'married'	''	''	
2	1	'female'	'3'	'YES'	''	'married'	'Univers	'working'	High Sch
3	2	'male'	'3'	'NO'	'Husband	'married'	''	''	
4	2	'female'	'3'	'YES'	''	'married'	'Univers	'working'	High Sch
5	3	'male'	'3'	'NO'	'Husband	'married'	''	''	
6	3	'female'	'3'	'YES'	''	'married'	'Univers	'working'	High Sch
7	4	'male'	'3'	'NO'	'Husband	'married'	''	''	
8	4	'female'	'3'	'YES'	''	'married'	'Univers	'working'	High Sch
9	5	'female'	'2'	'YES'	''	'never marri	''	'student'	
10	6	'male'	'4'	'YES'	''	'married'	'Univers	'working'	Dentist
11	6	'female'	'2'	'NO'	'Flat-ma	'never marri	''	'student'	
12	6	'male'	'2'	'NO'	'Flat-ma	'never marri	''	'student'	
13	6	'female'	'4'	'NO'	'Husband	'married'	'Primary	'Family-duties'	
14	7	'male'	'1'	'NO'	'Grandso	'never marri	''	''	
15	7	'male'	'7'	'YES'	''	'married'	'Upper s	'retired'	

Table 17: Excerpt from Individual table

Household ID	Block Code	Number of Hhld Members	Original Sample Number	Dwelling Ownership	Dwelling ID
1	16	2	70	Y	145
2	464	2	70	Y	9878
3	501	2	70	Y	11612
4	9999	2	70	Y	16840
5	47	1	69	Y	899
6	475	4	9	Y	10392
7	162	3	65	Y	3005
8	8070	1	20	N	16744
9	91	3	25	N	1674
10	248	3	25	N	4853
11	365	1	20	N	6807
12	552	1	20	N	13387
13	318	2	5	Y	5912
14	418	2	5	Y	8441
15	183	2	77	Y	3489

Table 18: Excerpt from Household table

Dwelling ID	Block Code	HHldNum	Building ID	Building Type	Number of Rooms
30	7	vacant	83	1	6
31	7	10273	150	1	4
32	7	9889	167	1	7
33	8	vacant	8897	1	7
34	8	vacant	9912	1	6
35	8	vacant	9911	1	5
36	8	vacant	9913	1	4
37	8	vacant	9907	1	6
38	8	vacant	9905	1	4
39	8	9981	8880	1	4
40	8	10862	9908	1	6
41	8	6806	9909	1	4
42	8	7043	9903	1	7
43	8	12007	9914	1	6
44	8	6425	8913	1	7
45	8	3627	9906	1	5

Table 19: Excerpt from Dwelling table

Building ID	Number of Floors	Block Code	Private Hhlds in block	Members in block	Dwellings in block	Number of Dwellings	Type
1	0	212	12	31	15	1	1
2	0	212	12	31	15	1	1
3	0	212	12	31	15	1	1
4	0	212	12	31	15	1	1
5	0	212	12	31	15	1	1
6	0	212	12	31	15	1	1
7	0	212	12	31	15	1	1
8	1	212	12	31	15	4	4
9	0	25	17	36	26	1	1
10	0	60	6	16	6	1	1
11	0	25	17	36	26	1	1
12	0	25	17	36	26	1	1
13	0	25	17	36	26	1	1
14	1	25	17	36	26	3	3
15	0	25	17	36	26	1	1

Table 20: Excerpt from Building table

Sub Area	Block Code	Block Code Num	Private Hhlds	Members	Dwellings
1,1	0000	0	0	0	22
1,1	0002	2	3	4	4
1,1	0003	3	0	0	0
1,1	0005	5	0	0	1
1,1	0007	7	2	5	5
1,1	0008	8	11	33	17
1,1	0009	9	0	0	0
1,1	0010	10	9	22	16
1,1	0011	11	5	9	8
1,1	0012	12	5	17	12
1,1	0013	13	9	30	12
1,1	0014	14	19	46	20
1,1	0015	15	6	12	11
1,1	0016	16	23	58	31
1,1	0017	17	7	17	13
1,1	0018	18	13	41	23
1,1	0019	19	15	37	24
1,1	0020	20	12	36	13

Table 21: Excerpt from Building block table

6.4 Probabilities

All parameters in *URM-Microsim* are probability distributions: what is the probability of a certain event occurring before January 1 of the next year. To determine whether a certain event will occur, the *Monte Carlo sampling method* is used. A random number between 0 and 1 is drawn from a uniform distribution, and is compared to the event probability (Oskamp, 1995). If the random number falls in the interval between 0 and the event probability then the event occurs; if the random number falls in the interval between the event probability and 1, the event does not occur.

In *URM-Microsim* each individual of the population is *sequentially* subjected to the probability of experiencing each *demographic* event (*death*, *fertility, union formation, union dissolution, flat-mate formation, flat-mate leaving* and *nest leaving*). Since events are simulated sequentially, the order of events is of importance mainly for two reasons (Oskamp, 1995). First the sequence may bias the number of events that occur. For example, if death is simulated before fertility, women who died may have had a child before they died. If fertility is simulated before death, births are included of women who would have died before they gave birth. Second the occurrence of one event may change the probability of occurrence of another event. For example, if marriage (union formation) is simulated first, then fertility probability increases since it depends on marital status. One strategy used in event ordering is the *random dates* of event occurring (Oskamp, 1995). This is a two-step process. The first step only determines if an event will occur, and if so on which of the 365 days of the year. The second step then executes the events in the order determined in the first step. In *URM-Microsim* it is assumed that an individual may experience only one event in one simulation period and therefore the second step executes the event that occurs earlier in the year.

The *probability distributions* required by *URM-Microsim* are calculated from the European Statistics (EUROSTAT) data tables based on 1997-2001 data for North Aegean, Lesvos Island or Municipality of Mytilene (http://epp.eurostat.ec.europa.eu).

6.4.1 Death event

Mortality is simulated in *URM-Microsim* using the death probability by *sex* and *age* calculated from the EUROSTAT data tables based on 1997-2001 data for North Aegean as shown in Table 22 and Figure 47.

Age specific death probability for national & local data 1997-2001												
Age in completed years	**Males Greece**			**Females Greece**			**Males North Aegean**			**Females North Aegean**		
	Population	**Deaths**	**Probability**	**Population**	**Deaths**	**Probability**	**Population**	**Deaths**	**Probability**	**Population**	**Deaths**	**Probability**
Less than 1	269.483	1.730	0,0064	254.288	1.352	0,0053	4.644	24	**0,0052**	4.524	28	**0,0062**
1	258.424	93	0,0004	243.736	84	0,0003	4.549	2	**0,0004**	4.228	3	**0,0007**
2	259.747	74	0,0003	244.563	48	0,0002	4.662	1	**0,0002**	4.254	1	**0,0002**
3	262.066	59	0,0002	247.136	40	0,0002	4.724	3	**0,0006**	4.276	0	**0,0000**
4	265.069	51	0,0002	249.091	50	0,0002	4.699	1	**0,0002**	4.347	0	**0,0000**
5 - 9	1.372.586	224	0,0002	1.294.701	166	0,0001	24.798	4	**0,0002**	22.752	2	**0,0001**
10 - 14	1.551.855	295	0,0002	1.460.470	198	0,0001	28.027	5	**0,0002**	25.826	3	**0,0001**
15 - 19	1.874.425	1.393	0,0007	1.758.164	487	0,0003	34.824	17	**0,0005**	29.641	8	**0,0003**
20 - 24	2.034.019	2.424	0,0012	1.935.774	627	0,0003	31.712	37	**0,0012**	30.156	6	**0,0002**
25 - 29	2.073.204	2.428	0,0012	2.006.672	705	0,0004	28.577	28	**0,0010**	29.467	8	**0,0003**
30 - 34	2.033.859	2.478	0,0012	2.015.074	906	0,0004	31.191	33	**0,0011**	30.695	16	**0,0005**
35 - 39	1.887.235	2.799	0,0015	1.899.868	1.199	0,0006	30.353	33	**0,0011**	29.724	24	**0,0008**
40 - 44	1.819.853	3.971	0,0022	1.831.677	1.810	0,0010	29.574	59	**0,0020**	28.022	20	**0,0007**
45 - 49	1.713.490	5.921	0,0035	1.708.903	2.669	0,0016	28.217	81	**0,0029**	27.228	46	**0,0017**
50 - 54	1.600.228	8.969	0,0056	1.632.641	3.861	0,0024	27.854	144	**0,0052**	28.439	69	**0,0024**
55 - 59	1.423.218	11.578	0,0081	1.503.671	5.087	0,0034	23.612	154	**0,0065**	25.386	72	**0,0028**
60 - 64	1.500.931	19.413	0,0129	1.655.089	9.232	0,0056	26.871	328	**0,0122**	29.585	154	**0,0052**
65 - 69	1.413.216	29.450	0,0208	1.592.065	16.227	0,0102	29.502	619	**0,0210**	32.167	302	**0,0094**
70 - 74	1.113.498	39.014	0,0350	1.356.202	27.424	0,0202	26.499	885	**0,0334**	30.804	508	**0,0165**

Age specific death probability for national & local data 1997-2001												
Age in completed years	**Males Greece**			**Females Greece**			**Males North Aegean**			**Females North Aegean**		
	Population	**Deaths**	**Probability**	**Population**	**Deaths**	**Probability**	**Population**	**Deaths**	**Probability**	**Population**	**Deaths**	**Probability**
75 - 79	680.381	38.554	0,0567	900.296	35.566	0,0395	18.352	1.025	**0,0559**	22.172	810	**0,0365**
80 - 84	398.117	39.132	0,0983	569.761	47.725	0,0838	11.561	1.151	**0,0996**	15.377	1.159	**0,0754**
85 and over	346.992	59.655	0,1719	495.992	72.000	0,1452	10.746	1.969	**0,1832**	13.842	2.305	**0,1665**

Table 22: Age specific death probability for national & local data

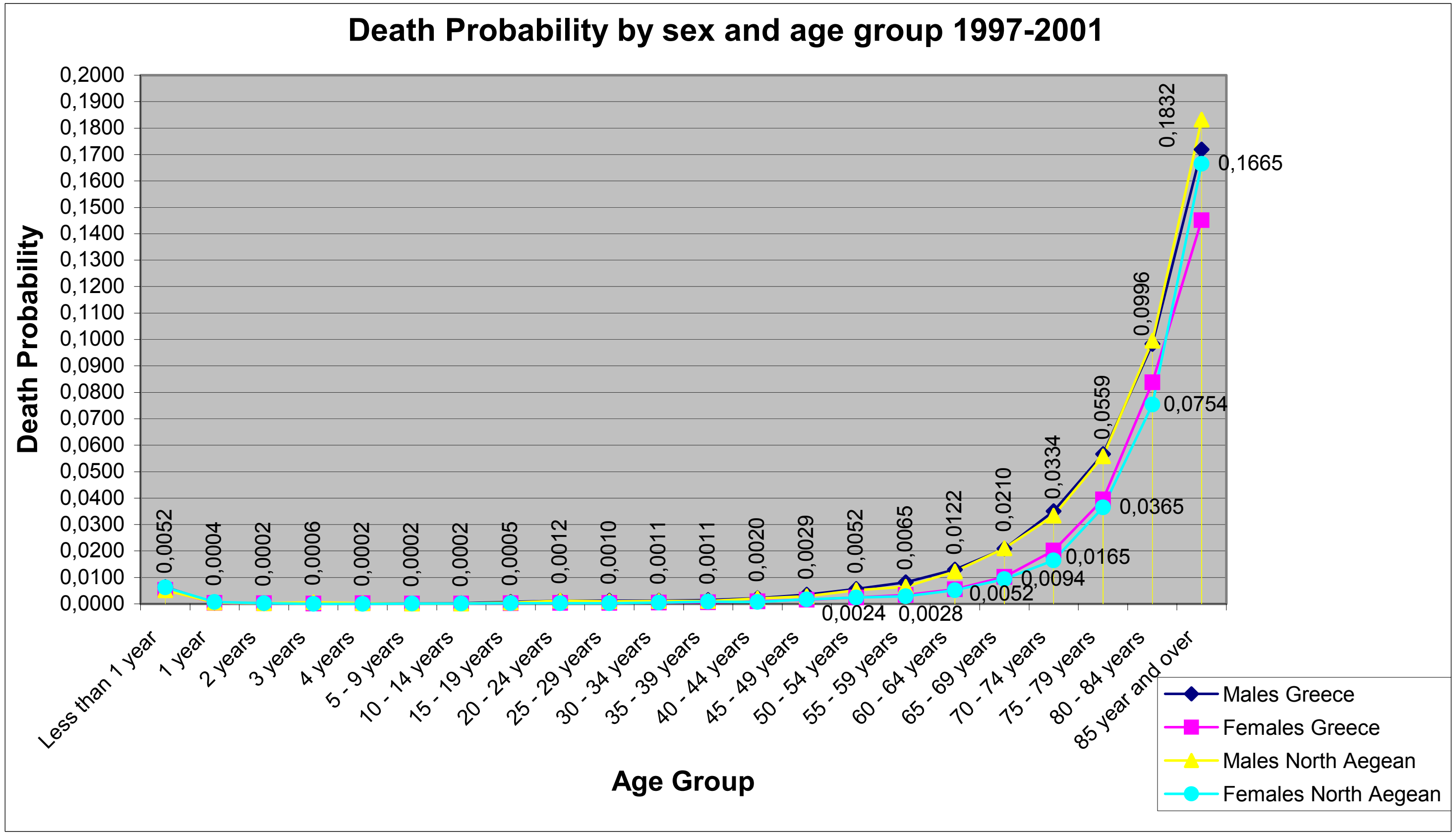


Figure 47: Death Probability by sex and age group 1997-2001

6.4.2 Fertility event

For the simulation of childbirth, four sets of probabilities are used: fertility probabilities, sex ratio at birth, mortality probability of newly born and twin probability. Within the context of this work *age specific* fertility probabilities as well as *age and marital status specific* probabilities for *national data* are calculated from the EUROSTAT data tables based on 1997-2001 data. In addition, *age specific* fertility probabilities for *local data (North Aegean)* are calculated from the EUROSTAT data tables based on 1997-2001 data. A comparison between national and North Aegean data shows that fertility levels in North Aegean are somewhat higher, and that women tend to have their children a little earlier (Figure 48). *Age and marital status specific* probabilities are not available for *local data* and therefore *national data* are used with corrections to deal with both the difference in tempo and quantum. The probability that a newly born child is female is **0,4935** for North Aegean (Table 25 and Figure 50). The mortality probability of newly born (that is, the probability that a new-born child will die in the same calendar year) is derived from the first row of Table 22 (that is, the probability that a newly born child will die under 1 year old) and is shown in Figure 51. Finally, the probability of a twin birth is **0.1** namely, the general probability found in researches in the Internet.

Age specific fertility probability for national & local data 1997-2001						
Age in completed years	Females Greece			Females North Aegean		
	Population	Births	Probability	Population	Births	Probability
10 - 14	1.460.470	329	0,0002	25.826	2	0,0001
15 - 19	1.758.164	19.119	0,0109	29.641	400	0,0135
20 - 24	1.935.774	97.120	0,0502	30.156	2.117	0,0702
25 - 29	2.006.672	176.448	0,0879	29.467	3.116	0,1057
30 - 34	2.015.074	149.874	0,0744	30.695	2.187	0,0712
35 - 39	1.899.868	55.152	0,0290	29.724	738	0,0248
40 - 44	1.831.677	9.954	0,0054	28.022	144	0,0051
45 - 49	1.708.903	917	0,0005	27.228	11	0,0004
50 and over	3.136.312	211	0,0001	53.825	0	0,0000

Table 23: Age specific fertility probability for national & local data

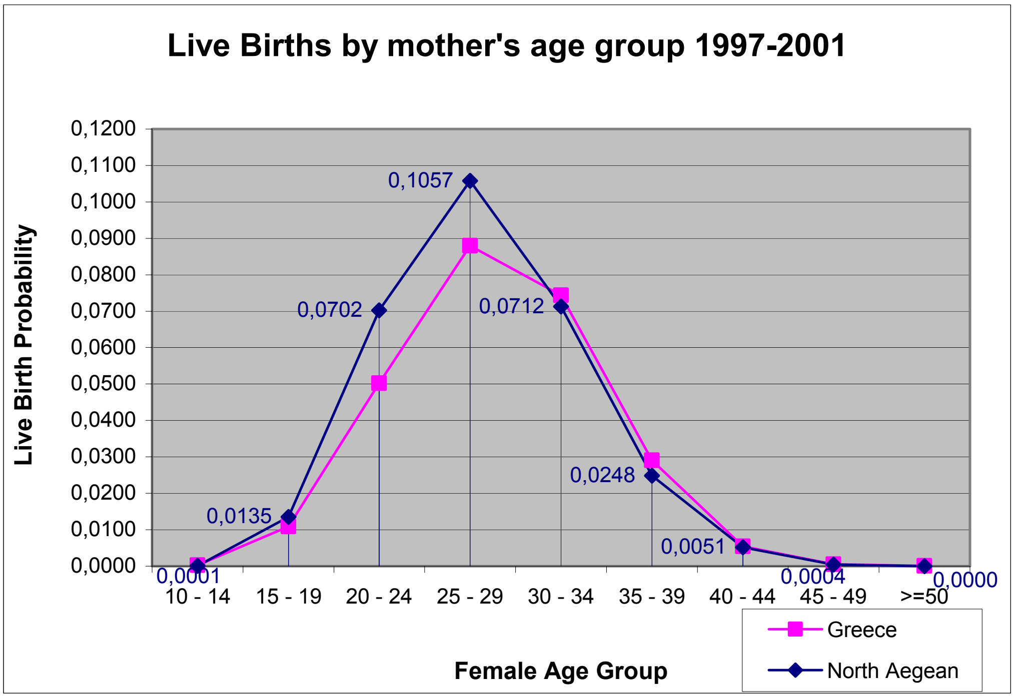


Figure 48: Live Births by mother's age group 1997-2001

Age & marital status specific fertility probability for national & local data 1997-2001								
Age in completed years	Females Greece						Females North Aegean	
	Married Females	In marriage Births	Prob	Not Married Females	Outside marriage Births	Prob	In marriage Prob	Outside marriage Prob
10 - 14	584	132	0,22595	1.459.886	197	0,00013	**0,07768**	**0,00005**
15 - 19	39.130	15.213	0,38878	1.719.034	3.906	0,00227	**0,48247**	**0,00282**
20 - 24	342.245	91.690	0,26791	1.593.529	5.430	0,00341	**0,37487**	**0,00477**
25 - 29	1.033.153	172.105	0,16658	973.519	4.343	0,00446	**0,20033**	**0,00536**
30 - 34	1.535.683	146.646	0,09549	479.391	3.228	0,00673	**0,09148**	**0,00645**
35 - 39	1.605.357	53.169	0,03312	294.511	1.983	0,00673	**0,02833**	**0,00576**
40 - 44	1.580.618	9.337	0,00591	251.059	617	0,00246	**0,00559**	**0,00232**
45 - 49	1.468.760	871	0,00059	240.143	46	0,00019	**0,00045**	**0,00014**
>=50	2.587.583	197	0,00008	548.729	14	0,00003	**0,00000**	**0,00000**

Table 24: Age & marital status specific fertility probability for national & local data

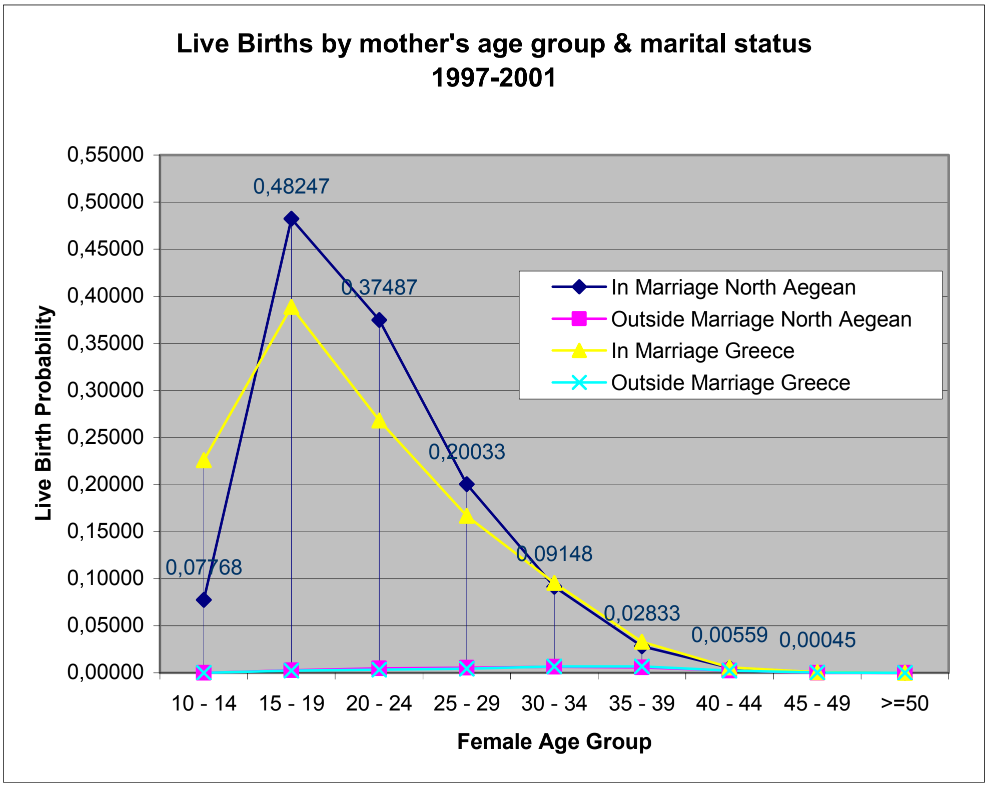


Figure 49: Live Births by mother's age group & marital status - N. Aegean 1997-2001

Living Newly Born Sex Probability						
Age in completed years 1997-1998	Population Greece			Population North Aegean		
	Total	Males	Females	Total	Males	Females
Less than 1 1997	100.043	51.134	48.909	1.819	904	915
Less than 1 1998	101.440	52.281	49.159	1.753	967	786
Less than 1 1999	100.241	51.629	48.612	1.706	856	850
Less than 1 2000	100.024	51.542	48.482	1.734	861	873
Less than 1 2001	122.023	62.897	59.126	2.156	1.056	1.100
Less than 1 Total	523.771	269.483	254.288	9.168	4.644	4.524
Sex Probability		**0,5145**	**0,4855**		**0,5065**	**0,4935**

Table 25: Living Newly Born Sex Probability

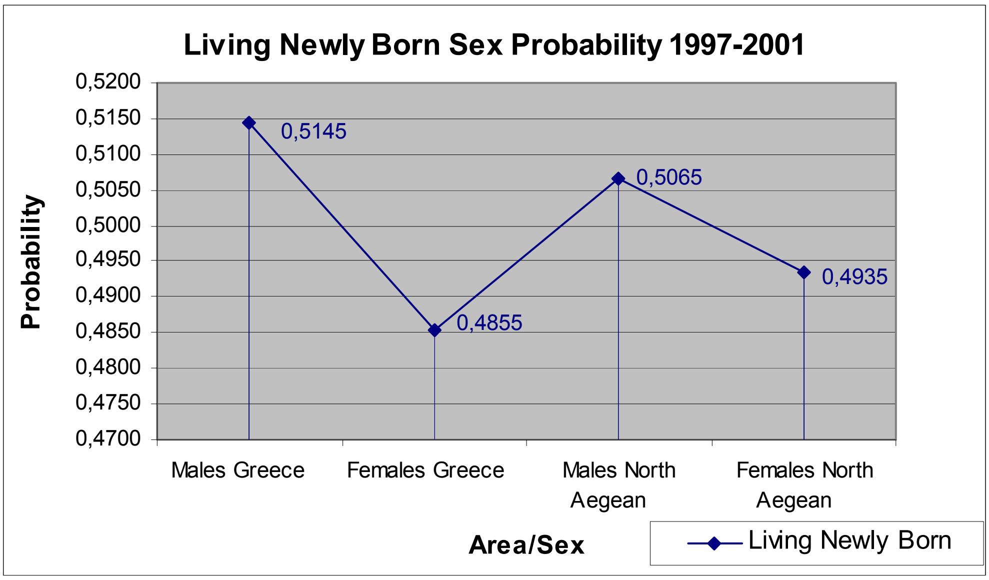


Figure 50: Living Newly Born Sex Probability 1997-2001

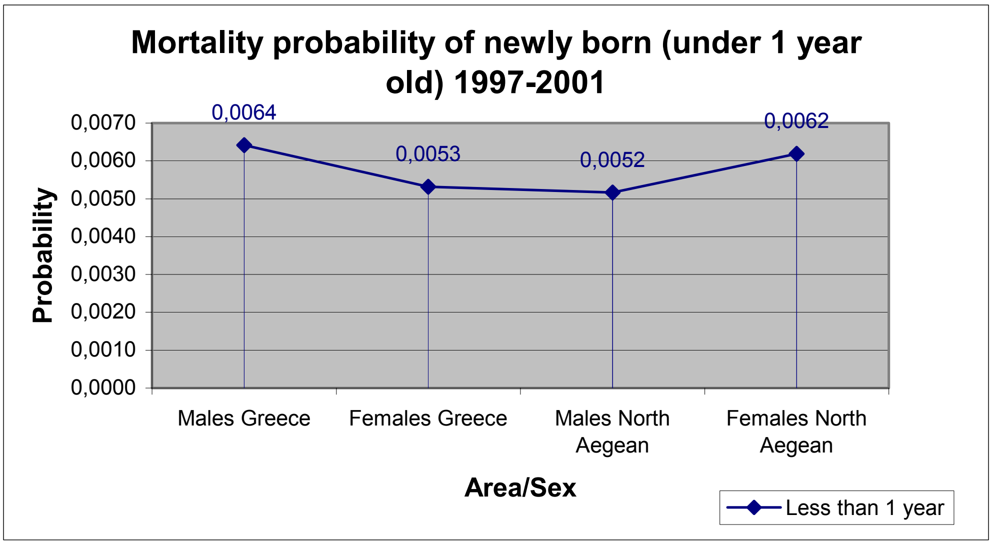


Figure 51: Mortality probability of newly born (under 1 year old) 1997-2001

6.4.3 Union formation event

In *URM-Microsim* three types of union formation are distinguished: *marriage*, *cohabitation* and *marriage of cohabitating partners*. The latter is, technically, not a union formation since no partner search takes place, just a change in partners' marital status.

Union formation probabilities are not available for Greece and therefore *marriage* probabilities are used. This is not a large misspecification since in Mytilene most partners do not cohabitate: they live in separate households until they get married. If partners cohabitate then either at least one of them comes for another municipality (in-migration) or they get married shortly after cohabitation. It should be noted that the questionnaire data analysis points out that each year roughly one (1) household of cohabiters is formed out of 109 households. The *cohabitation* probability is calculated from the questionnaire data analysis and is shown in Table 29.

Sex and *age specific marriage* probabilities are calculated from the NSS data tables based on 1998-2001 data for Greece (Table 26). In addition, the probabilities *male of age x marrying female of age y* as well as the probabilities *female of age x marrying male of age y* are calculated from the NSS data tables based on 1998-2001 data for Greece as shown in Table 27 and Table 28 respectively. These probabilities are calculated from marriages that actually took place. Marriages that were planned but could not be effectuated due to lack of housing opportunities are not taken into consideration. This implies that the model needs intension marriage probabilities. Since these are not available, probabilities based on actual behavior are used.

If two partners forming a union are both singles that is they both have their own dwellings then a decision must be made on whose dwelling to keep. As no statistical data exist, a common sense rule is implemented in *URM-Microsim*. If one of the partners is owner-occupant then the new couple will keep the owner-occupied dwelling. If both partners are renters or owner-occupants then they will keep the larger of the two dwellings. In case of two dwellings of the same size, the partners will keep the female's dwelling. If only one of the partners has a dwelling then the couple will naturally live there. If neither partner has a dwelling of his/her own, then a new dwelling has to be found.

In case of *marriage of cohabitating partners* no statistics exist, and the probability that a cohabitating couple decides to get married is assumed to be **0.5.**

Marriage Probability by age of unmarried population Greece 1998-2001						
Age Group	Males Greece			Females Greece		
	Not Married	Marriages	Probability	Not Married	Marriages	Probability
< 15	1.222.162	18	0,0000	1.148.143	301	0,0003
15-19	1.487.701	1.731	0,0012	1.361.684	13.703	0,0101
20-24	1.632.466	22.215	0,0136	1.275.427	61.184	0,0480
25-29	1.663.924	76.342	0,0459	780.155	82.106	0,1052
30-34	1.646.297	66.891	0,0406	387.214	39.253	0,1014
35-39	1.517.618	27.836	0,0183	236.735	13.410	0,0566
40-44	1.468.135	12.490	0,0085	202.919	6.103	0,0301
45-49	1.379.467	6.113	0,0044	193.399	3.376	0,0175
50-54	1.294.397	3.777	0,0029	213.850	2.178	0,0102
55-59	1.127.836	2.156	0,0019	223.845	1.047	0,0047
60-64	1.195.398	1.664	0,0014	340.768	745	0,0022
65-69	1.136.667	1.290	0,0011	445.382	371	0,0008
70-74	912.629	880	0,0010	536.509	184	0,0003
>=75	1.148.950	622	0,0005	1.160.974	64	0,0001

Table 26: Marriage Probability by age of unmarried population Greece 1998-2001

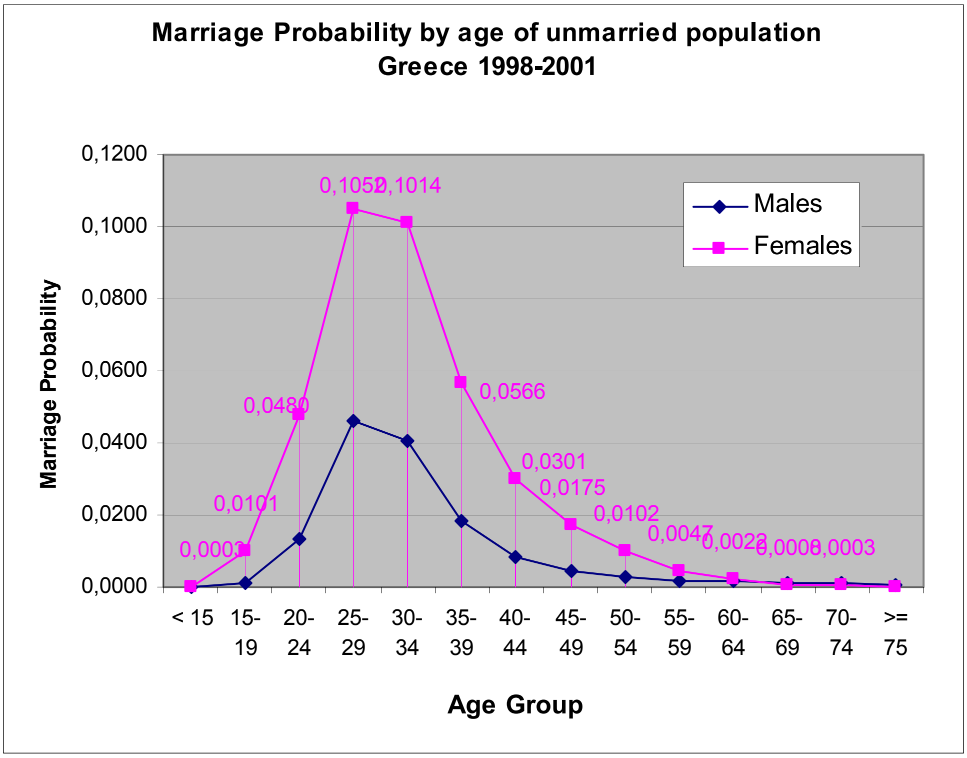


Figure 52: Marriage Probability by age of unmarried population Greece 1998-2001

Probability Male of age x marrying Female of age y Greece 1998-2001													
Male Age Group	Female Age Group												
	< 15 years	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	>=70
< 15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
15-19	0,0001	0,0008	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
20-24	0,0001	0,0036	0,0077	0,0018	0,0003	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
25-29	0,0000	0,0028	0,0178	0,0212	0,0034	0,0005	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
30-34	0,0000	0,0009	0,0087	0,0187	0,0104	0,0016	0,0003	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000
35-39	0,0000	0,0002	0,0021	0,0062	0,0064	0,0027	0,0006	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000
40-44	0,0000	0,0000	0,0005	0,0018	0,0027	0,0021	0,0010	0,0003	0,0001	0,0000	0,0000	0,0000	0,0000
45-49	0,0000	0,0000	0,0001	0,0006	0,0010	0,0011	0,0010	0,0005	0,0001	0,0000	0,0000	0,0000	0,0000
50-54	0,0000	0,0000	0,0001	0,0002	0,0004	0,0006	0,0007	0,0006	0,0003	0,0001	0,0000	0,0000	0,0000
55-59	0,0000	0,0000	0,0000	0,0001	0,0001	0,0002	0,0004	0,0004	0,0004	0,0002	0,0001	0,0000	0,0000
60-64	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0002	0,0003	0,0003	0,0002	0,0001	0,0000	0,0000
65-69	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0002	0,0003	0,0002	0,0002	0,0001	0,0000
>=70	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0002	0,0002	0,0002	0,0001	0,0001

Table 27: Probability Male of age x marrying Female of age y Greece 1998-2001

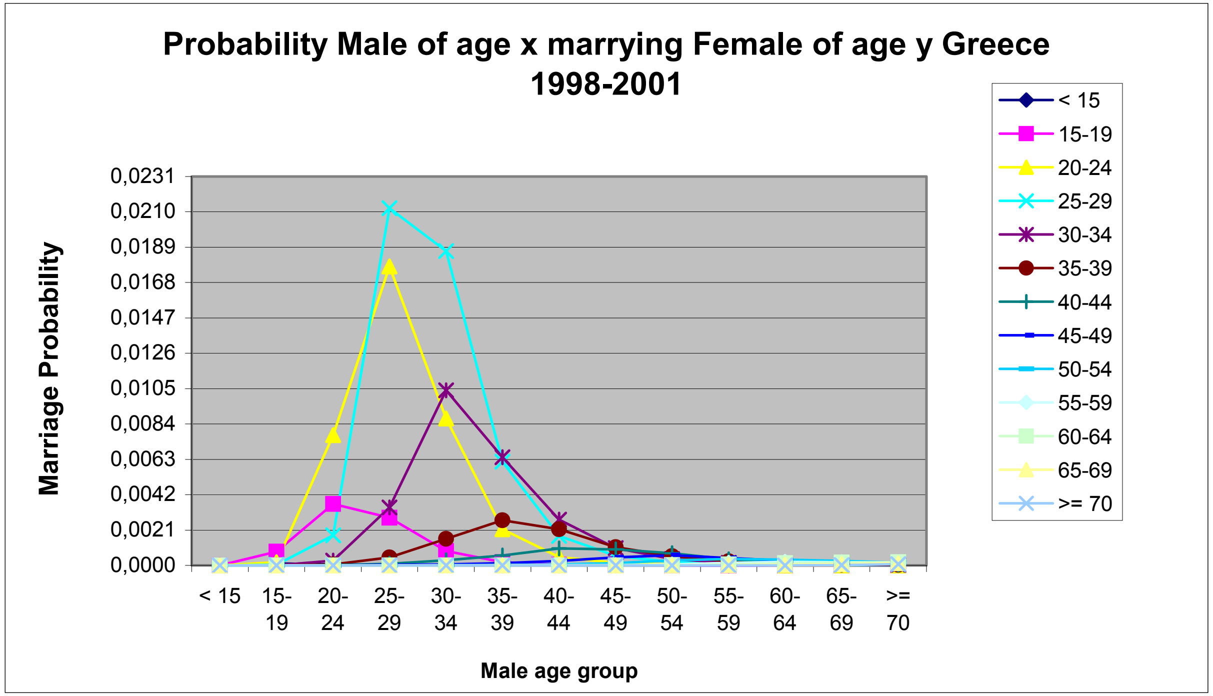


Figure 53: Probability Male of age x marrying Female of age y Greece 1998-2001

Probability Female of age x marrying Male of age y Greece 1998-2001													
Male Age Group	**Female Age Group**												
	< 15 years	**15-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50-54**	**55-59**	**60-64**	**65-69**	**>= 70**
< 15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
15-19	**0,0001**	**0,0009**	**0,0002**	**0,0001**	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
20-24	**0,0001**	**0,0044**	**0,0099**	**0,0038**	**0,0012**	**0,0004**	**0,0001**	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
25-29	0,0000	**0,0035**	**0,0232**	**0,0452**	**0,0148**	**0,0033**	**0,0007**	**0,0003**	**0,0001**	0,0000	0,0000	0,0000	0,0000
30-34	0,0000	**0,0010**	**0,0113**	**0,0394**	**0,0442**	**0,0109**	**0,0024**	**0,0005**	**0,0002**	0,0000	0,0000	0,0000	0,0000
35-39	0,0000	**0,0002**	**0,0026**	**0,0120**	**0,0252**	**0,0171**	**0,0044**	**0,0010**	**0,0002**	**0,0001**	0,0000	0,0000	0,0000
40-44	0,0000	**0,0001**	**0,0006**	**0,0033**	**0,0103**	**0,0133**	**0,0073**	**0,0020**	**0,0004**	**0,0001**	0,0000	0,0000	0,0000
45-49	0,0000	0,0000	**0,0002**	**0,0010**	**0,0036**	**0,0064**	**0,0065**	**0,0033**	**0,0009**	**0,0001**	0,0000	0,0000	0,0000
50-54	0,0000	0,0000	**0,0001**	**0,0003**	**0,0012**	**0,0030**	**0,0046**	**0,0041**	**0,0019**	**0,0003**	**0,0001**	0,0000	0,0000
55-59	0,0000	0,0000	0,0000	**0,0001**	**0,0004**	**0,0011**	**0,0020**	**0,0026**	**0,0020**	**0,0008**	**0,0002**	0,0000	0,0000
60-64	0,0000	0,0000	0,0000	0,0000	**0,0002**	**0,0005**	**0,0011**	**0,0018**	**0,0019**	**0,0011**	**0,0005**	**0,0001**	0,0000
65-69	0,0000	0,0000	0,0000	0,0000	**0,0001**	**0,0003**	**0,0006**	**0,0010**	**0,0014**	**0,0010**	**0,0006**	**0,0002**	0,0000
>=70	0,0000	0,0000	0,0000	0,0000	0,0000	**0,0001**	**0,0002**	**0,0006**	**0,0007**	**0,0006**	**0,0005**	**0,0003**	**0,0001**

Table 28: Probability Female of age x marrying Male of age y Greece 1998-2001

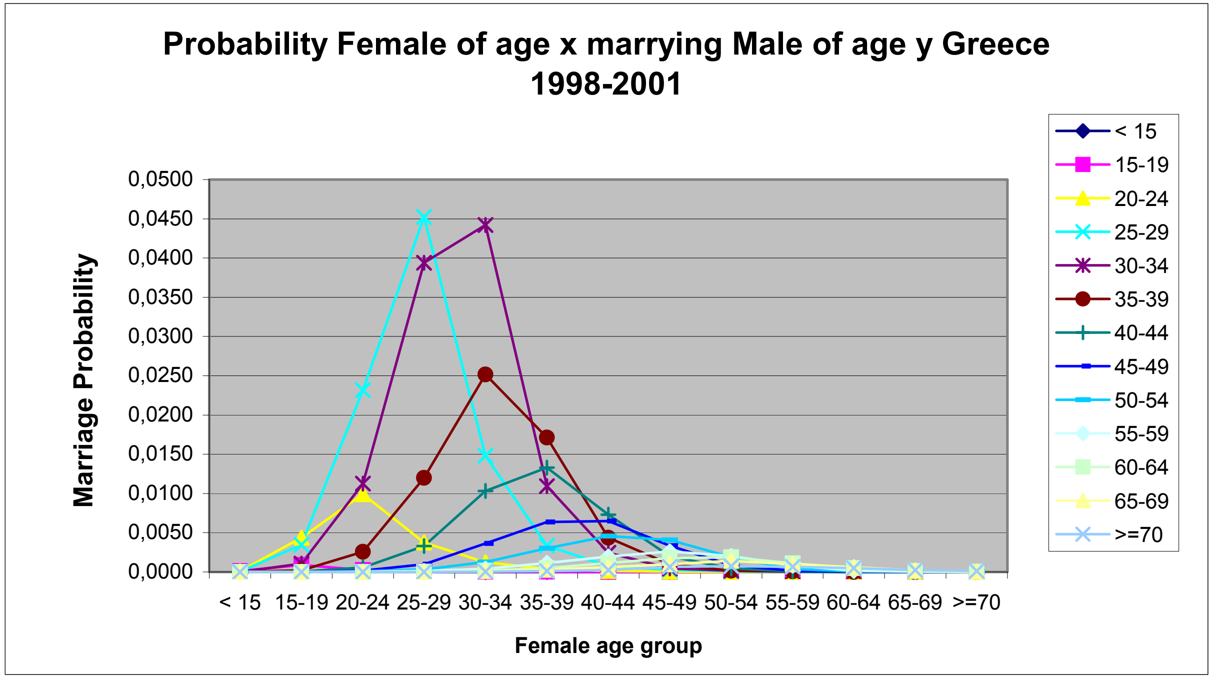


Figure 54: Probability Female of age x marrying Male of age y Greece 1998-2001

Cohabitate Probability Mytilene Questionnaire 2005	
Total Population 2005	280
Not Married Population 2005	166
Mean Number of Cohabitates 2000-2005	1,6667
Cohabitate Probability	**0,0100**

Table 29: Cohabitate Probability Mytilene 2005

6.4.4 Union dissolution event

Table 30 and Figure 55 show the *divorce rates* EUROSTAT data table by *duration of marriage* based on 1997-1999 data for Greece. For each calendar year n, if the number of divorces ranked according to the duration of marriage in years x is available, divorce rates by duration of marriage can be calculated by relating the number of divorces at the end of x years of marriage to the number of marriages in year n-x. *Age specific marriage duration* probabilities are calculated from the NSS data tables based on 2001 data for Mytilene as shown in Table 31 and Figure 56. The age is the age of the female partner of marriage. For each married individual, the duration of his/her marriage is calculated based on the probabilities of Table 31. Next, the individual is subjected to the probability of divorcing based on his/her years of marriage (Table 30).

In case of *de-habitation* no statistics exist, and the probability of union dissolution of cohabiters is **0.1** namely, the general probability found in researches in the Internet.

In *URM-Microsim* union dissolution is subject to housing opportunities: a union cannot dissolve if there is no available dwelling for the leaving partner. This implies that the model needs intension union dissolution probabilities. Since these are not available, probabilities based on actual behavior are used. In the case of union dissolution this is not a misspecification since it is not very likely, that partners who want their union to dissolve will decide against it because no housing is available. It is more likely that the leaving partner will move in with his/her parents, occupy a dwelling, which is not according to his/her needs, or move out of the municipality. It should be noted that the questionnaire data analysis points out that 3 out of 5 divorced individuals are living with their parents.

As a union dissolves, it should be decided whether the female or the male moves out of the dwelling and who keeps the children. In Greece custody is granted to the mother. Therefore, the children stay with the mother and any singles in the household are assumed to be the mother's parents. In *URM-Microsim* the probability that the *male moves out* is **0.85** since the dwelling in which a couple lives usually belongs to the female (there is a custom in Mytilene according to which a bride should have a house) and even if this is not the case, a male allows his partner and children to keep the house.

The leaving partner is subjected to the probability of out-migrating. The partner is highly probable to out-migrate if he/she is a citizen of another Municipality. In Table 32 the *divorced out-migration probability* is calculated from the NSS data tables based on 2001 data. It should be noted that the difference between divorced females and males could be owed to the fact that divorced males have got re-married or the fact that divorced females that lived in another municipality returned to the parental municipality (in-migration) after the divorce. However, since these are the only data available it is decided to use them in order to have a rough estimate of the divorced out-migrate probability.

Divorce Rates By Duration Of Marriage Greece 1997-1999 EUROSTAT data table									
Year	Duration of marriage reached during the year								
	Less than 1	1	2 to 4	5 to 9	10 to 14	15 to 19	20 to 24	25 to 29	30 or more
1997	0,0000	0,0023	0,0092	0,0085	0,0054	0,0040	0,0025	0,0018	0,0083
1998	0,0000	0,0014	0,0084	0,0076	0,0051	0,0032			
1999	0,0000	0,0016	0,0106	0,0093	0,0060	0,0041	0,0026	0,0028	
Total	**0,0000**	**0,0018**	**0,0094**	**0,0085**	**0,0055**	**0,0038**	**0,0026**	**0,0023**	**0,0083**

Table 30: Divorce Rates By Duration Of Marriage Greece 1997-1999

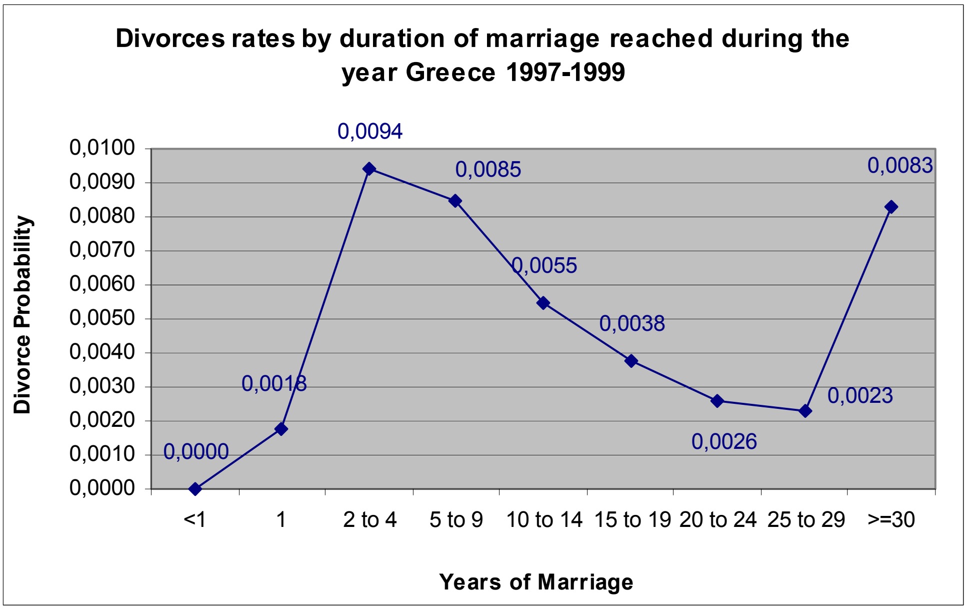


Figure 55: Divorces rates by duration of marriage

Female's age specific Marriage Duration probability Mytilene 2001									
Marriage Duration	Female's Age in completed years								
	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55+
1 year	0,0457	0,2690	0,3147	0,1726	0,0508	0,0305	0,0152	0,0152	0,0863
2 to 4	0,0238	0,1997	0,3592	0,2191	0,0760	0,0358	0,0119	0,0134	0,0611
5 to 9	0,0020	0,0602	0,2929	0,3320	0,1494	0,0652	0,0241	0,0160	0,0582
10 to 14	0,0000	0,0020	0,0897	0,3675	0,2611	0,1143	0,0414	0,0305	0,0936
15 to 19	0,0000	0,0000	0,0043	0,1298	0,3564	0,2383	0,1074	0,0468	0,1170
20 to 24	0,0000	0,0000	0,0000	0,0049	0,1430	0,3619	0,2298	0,0937	0,1667
25 to 29	0,0000	0,0000	0,0000	0,0000	0,0031	0,1454	0,3402	0,2443	0,2670
>=30	0,0000	0,0000	0,0000	0,0000	0,0000	0,0008	0,0359	0,1183	0,8450

Table 31: Female's age specific Marriage Duration probability Mytilene 2001

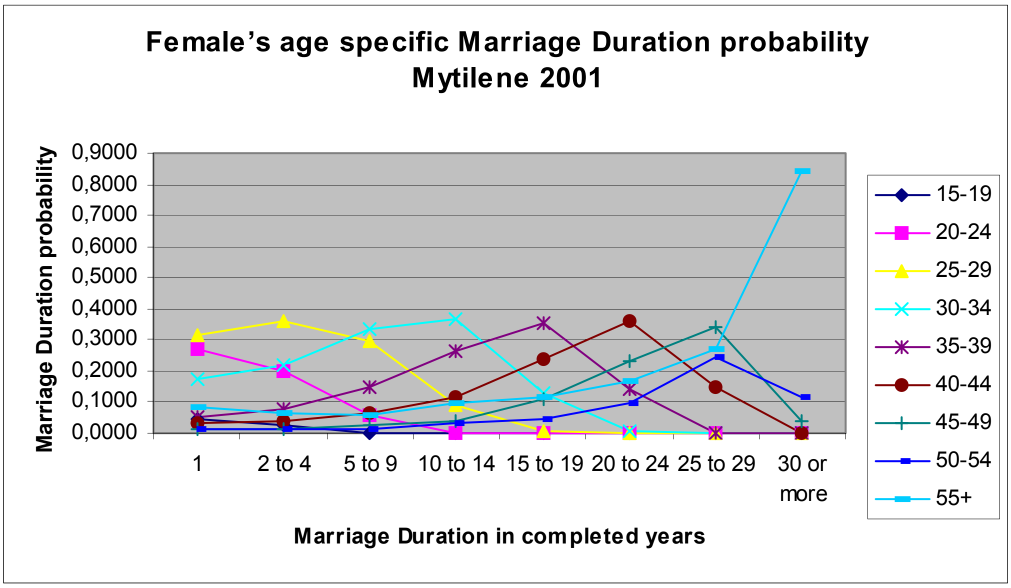


Figure 56: Female's age specific Marriage Duration probability Mytilene 2001

Divorced Out-Migrate Probability Prefecture of Lesvos 2001	
Population	**Divorced**
Total	2.705
Males	1.148
Females	1.557
The difference between divorced females & males	409
Divorced Out-Migrate Probability	**0.1516**

Table 32: Divorced Out-Migrate Probability Prefecture of Lesvos 2001

6.4.5 Nest leaving event

In *URM-Microsim* nest leaving is defined as the process of leaving the parental home to live alone. Children leaving the parental home to live together (cohabitation or marriage) are treated in the union formation event.

Nest leaver may choose to live inside Mytilene or outside. Therefore, each individual of age 18 and above, who still lives with his/her parents is initially subjected to the probability of out-migrating. According to the School Administrative Unit of Lesvos, 60% of school children of age 18 to 20 succeed in the Panellenic examinations and are accepted in University Departments most of which reside in other municipalities. Although some of these school children may be accepted in one of Mytilene's University Departments (Mytilene has six University Departments) or they may decide to remain in Mytilene and take part in the next year's Panellenic examinations, the majority of nest leavers for reasons of education will move out of Mytilene. In addition, there are a proportion of children of age 18 and above, who out-migrate for work-related reasons (for example to search for a job in Athens or to work in the Army). Children that choose to live outside reach this decision independently of the housing market opportunities within the municipality. The *child out-migrate* probability is **0.6** and is calculated from the School Administrative Unit of Lesvos.

The rest of the children of age 18 and above, either already work or once they graduate high school they search for a job in Mytilene. In Mytilene children do not often leave the parental home to live alone. They usually live with their parents until they get married. This also applies to children who out-migrated for educational reasons and returned to Mytilene once they completed their studies. The *nest leaving* probability is calculated from the questionnaire data analysis and is shown in Table 33. It should be noted that nest leaving is a process which highly depends on housing market opportunities: dependent children wishing to leave the parental home may only do so if housing is available. This implies that the model needs intension nest leaving probability. Since this is not available, the probability based on actual behavior is used.

Nest Leaving Probability Mytilene Questionnaire 2005	
Number of Children of age >= 18	38
Mean Number of Nest Leavers 2000-2005	0,6667
Nest Leaving Probability	**0,0175**

Table 33: Nest Leaving Probability Mytilene 2005

6.4.6 Flat-mate formation event

In *URM-Microsim* flat-mate formation is defined as the process during which an individual who lives on his/her own or flat-mates with others in a rented dwelling finds another individual of the same sex to flat-mate. Flat-mate formation is a common practice especially for students and Army officers.

The *flat-mate formation* probability is calculated from the questionnaire data analysis and is shown in Table 34. This probability is calculated from flat-mate formations that actually took place. Formations that were planned but could not be effectuated due to lack of housing opportunities are not taken into consideration. This implies that the model needs intension flat-mate formation probability. Since this is not available, the probability based on actual behavior is used.

If two individuals deciding to flat-mate both live on their own, that is, they both have their own dwellings, then a decision must be made on who's dwelling to keep. As no statistical data exist, a common sense rule is implemented in *URM-Microsim*. The two flat-mates will keep the larger of the two dwellings. In case of two dwellings of the same size, the flat-mates will keep the cheaper dwelling. If only one of the flat-mates lives on his/her own and therefore has a dwelling, then the flat-mates will naturally live there. If neither flat-mate has a dwelling of his/her own, then a new dwelling has to be found.

Flat-mate Formation & Leaving Probabilities Mytilene Questionnaire 2005		
Students/Military Workers etc Population	25	%
Flat-mate	8	32
Living alone	15	60
Number of flat-mate formation persons	2	
Flat-mate Formation Probability	**0,08**	
Number of flat-mate dissolution persons	2	
Flat-mate Dissolution Probability	**0,08**	

Table 34: Flat-mate Formation & Leaving Probabilities Mytilene 2005

6.4.7 Flat-mate leaving event

In *URM-Microsim* flat-mate leaving is defined as the process of leaving the flat-mate house to live alone. Flat-mates leaving the flat-mate house to live together are treated in the flat-mate formation event.

A flat-mate may choose to live inside the Mytilene or outside. Therefore, each flat-mate is initially subjected to the probability of out-migrating. It should be noted that most students leave Mytilene once they have completed their studies. Since no statistical data are available, it is assumed that the out-migration probability of a student of age 25 and above is **0.9**, while all other flat-mates' out-migration probability is **0.1**.

The *flat-mate leaving* probability is calculated from the questionnaire data analysis and is shown in Table 34. It should be noted that flat-mate leaving is a process, which highly depends on housing market opportunities: flat-mates wishing to leave the flat-mate house may only do so if housing is available. This implies that the model needs intension flat-mate leaving probability. Since this is not available, the probability based on actual behavior is used.

6.4.8 Out-migration event

In *URM-Microsim* all households that fail to find a dwelling (except in-migrating households) are subjected to the probability of out-migrating. The *out-migration* probability highly depends on the type of event experienced by the household. For example, the leaving partner of a union is more likely to out-migrate in another municipality in case of dwelling search failure, than the household whose size increases due to childbirth. Since no statistical data are available, event specific out-migration probabilities are assumed, taking into consideration the fact that households would substitute one or more of their search criteria to obtain a dwelling in the area of Mytilene, and are shown in Table 35.

Event	Out-Migration Probability
Total	0,4
Death	0,05
Fertility	0,05
Nest Leavers	0,01
Union dissolution Leaver	0,1
Union formation Leavers	0,05
Union formation (one dwelling)	0,01
Flat-mate formation Leavers	0,05
Flat-mate formation (one dwelling)	0,01
Flat-mate leaving Leaver	0,05
Stayers	0,01

Table 35: Event specific Out-migration probabilities

6.4.9 In-migration event

Within the housing market system of Mytielene*,* the most important categories of in-migrating households are:

1. *Students*. According to the information available in the site of the University of the Aegean (www.aegean.gr), each year about 700

students are accepted in the Departments of the University that are located in Mytilene. In addition, the probability that an in-migrating student is female is **0.51.** Moreover, according to the Mytilene Questionnaire 40% of students flat-mate mostly in pairs (at least when in-migrating), that is 280 students (140 households), and the remaining 60% of students live alone, that is 420 students (420 households). Therefore, it is assumed that each year about 560 households in-migrate.

2. *Military men.* Another category of in-migrating households is Military men and their families who are obliged by their job to move from town to town every two to five years. Since the Lesvos Island and Mytilene are at the bounds of Greece, a large number of military camps are located at the boundaries of the city.
3. *Foreigners.* About 2.000 foreigners, who mostly come from East Europe, live in Lesvos Island. A significant number of the above mentioned lives and works in the city of Mytilene.
4. *Urbanization.* Individuals and households move from other villages of the island to Mytilene, usually searching for better job opportunities.

Since the only statistics available concern student in-migration and it is estimated that it constitutes the largest in-migration category, the remaining categories were not taken into consideration.

6.4.10 Education event

In *URM-Microsim sex and age specific* education probabilities for *national data* are calculated from the EUROSTAT data tables based on 2001 data (Table 36 and Figure 57).

Age specific Education probability for national data 2001								
Highest level of educational attainment	Males – Age				Females -Age			
	15-24	25-34	35-54	55-74	15-24	25-34	35-54	55-74
1. No education at all	**0,0111**	**0,0134**	0,0123	0,0373	**0,0101**	**0,0110**	0,0150	0,1085
2. Primary or less	**0,1332**	**0,1527**	0,3271	0,6187	**0,0915**	**0,1253**	0,3931	0,6757
3. Lower secondary	**0,3109**	**0,1509**	0,1022	0,0578	**0,2834**	**0,1211**	0,0889	0,0369
4. Upper secondary	**0,4493**	**0,4117**	0,3101	0,1607	**0,4720**	**0,3929**	0,2963	0,1262
5. Post secondary	**0,0495**	**0,0569**	0,0344	0,0146	**0,0827**	**0,0885**	0,0368	0,0085
6. Tertiary	**0,0461**	**0,2144**	0,2138	0,1108	**0,0603**	**0,2612**	0,1698	0,0442

Table 36: Age specific Education probabilities Greece 2001

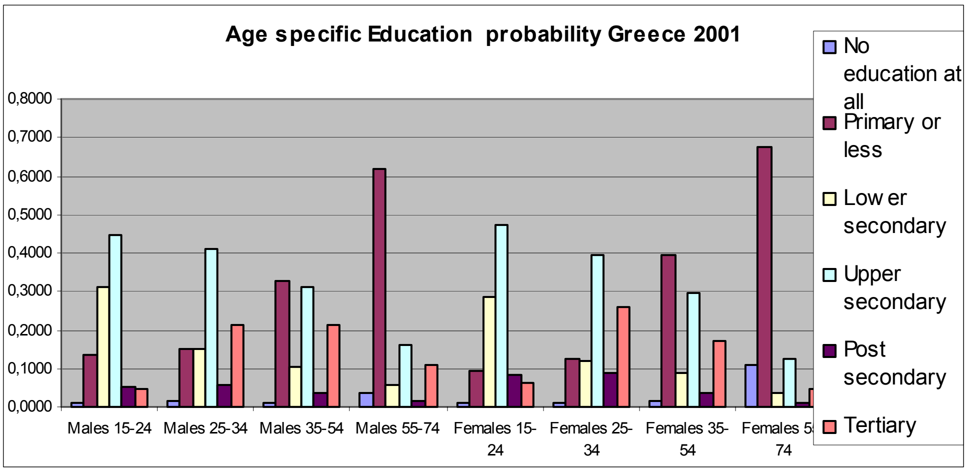


Figure 57: Age specific Education probability Greece 2001

6.4.11 Demand change event

In *URM-Microsim* all households that do not experience a demographic or income change event are subjected to the probability of demand changing. Since no statistical data are available and no realistic assumption can be made, a low demand change probability is assumed (**0.001)**.

6.4.12 Residential search and migration use case

Finding and securing a dwelling and being successful in housing search is a process governed by a number of probabilities. First there is the *acceptance probability*: will the searching household accept the dwelling offered? The acceptance probabilities are disaggregated by the four dimensions of the housing aspirations that are of importance (primary dimensions), namely *size (number of rooms)*, *price*, *type* and *tenure*. The desired dimensions of a dwelling highly depend on current household tenure (if any) and on the event experienced by the household. The probabilities that determine the desired dwelling are calculated from the Mytilene Questionnaire data and are shown in Table 38**.** As mentioned in **Chapter 5**, dwellings are defined by seven dimensions. However no statistics are available to determine the remaining three dimensions of a dwelling (secondary dimensions): *structure condition, amenities (garden, parking etc),* and *neighborhood quality*. Therefore, the probability that a household may reject an offered dwelling that has all desired primary dimensions but falls short of the secondary dimensions, should be taken into consideration. That is why the OwnerAcceptance and the RenterAcceptance probabilities are used namely the probability of an owner to accept the offered dwelling and the probability of a renter to accept the offered dwelling correspondingly. These probabilities are calculated from the Mytilene Questionnaire data and are shown in Table 37 and Table 38**.**

Aside from the probabilities discussed above, search intensity is an important factor governing the housing search process. As mentioned in **Chapter 4**, all households do not have an equally strong desire to find a new dwelling; the *search intensity* with which a household searches for a new dwelling depends on the event that triggers the desire to move and determines the number of dwellings the household will be offered. The number of dwelling offers highly depends on the housing market opportunities.

Ownership specific Residential Search probabilities Mytilene Questionnaire 2005	
Num of Households	109
Num of Households searching	16
Probability of Household searching	**0,1468**
Num of Owner Households searching	6
Probability of Owner Household searching (pOwnerAcceptance)	**0,375**
Num of Owner Households search for a dwelling to buy	4
PownerBuy	**0,67**
Num of Owner Households search for a dwelling to rent	2
PownerRent	**0,33**
Num of Renter Households searching	10
Probability of Renter Household searching (pRenterAcceptance)	**0,625**
Num of Renter Households search for a dwelling to buy	5
PrenterBuy	**0,5**
Num of Renter Households search for a dwelling to rent	5
PrenterRent	**0,5**

Table 37: Ownership specific Residential Search probabilities Mytilene 2005

Event and Search Criteria specific Residential Search probabilities Mytilene Questionnaire 2005

Event	PownerBuy	pOwnerRent	PPenterBuy	pRenterRent	POwnerAcceptance	pRenterAcceptance	pLarger	pCheaper
					0,375	0,625		
Death	0,67	0,33	0,2	0,8	0,06	0,06	0,01	0,5
Fertility	0,67	0,33	0,5	0,5	0,06	0,06	0,9	0,5
Nest Leavers	0	0	0,2	0,8	0	0,8	0,1	0,5
Stayers	0,67	0,33	0,5	0,5	0,015	0,03	0,4	0,5
Union dissolution Leaver	0	0	0,2	0,8	0	0,9	if mother & children leave then 0,5 else 0,1	0,5
Union formation Leavers	0	0	0,5	0,5	0	0,06	0,1	0,5
Union formation (one dwelling)	0,67	0,33	0,5	0,5	0,06	0,03	0,4	0,5
Flat-mate formation Leavers	0	0	0	1	0	0,06	0,1	0,5
Flat-mate formation (one dwelling)	0	0	0	1	0	0,03	0,6	0,5
Flat-mate leaving Leaver	0	0	0	1	0	0,06	0,1	0,5
Flat-mate Stayers	0	0	0	1	0	0,03	0,2	0,5
In-migration students	0	0	0	1	0	0,9	0,1	0,8
Demand change	0,67	0,33	0,5	0,5	0,18	0,205	0,5	0,5

Table 38: Event & Search Criteria specific Search probabilities Mytilene 2005

6.5 Summary

This chapter was devoted to applying *URM-Microsim* in the real world. The Municipality of Mytilene was chosen as a testing ground for *URM-Microsim*. Initially, a short introduction on the history and background of Mytilene, its population and its housing market were presented. Next, the population synthesis technique used to produce the Mytilene data was presented as well as its input data and an excerpt of the synthetic data is given. Finally, the probability distributions required by *URM-Microsim* are calculated.

Chapter 7
Results and Discussion

7.1 Introduction

This chapter focuses on the performance of *URM-Microsim*. Once the system is loaded with the Mytilene data, it is executed in order to simulate the population and housing market developments. Then, the issue of model efficiency arises: how well does the model simulate reality? Section 7.2 discusses the notion of system calibration, presents the series of simulations carried out for *URM-Microsim* and discusses the simulation results. Next, in section 7.3 the notion and methods of system validation are discussed.

7.2 Calibration

Once *URM-Microsim* is loaded with the Mytilene data for 2001, it is executed in order to simulate the population and housing market developments on a yearly basis. During program execution, a variety of errors may occur owning either to the inherent complexity of residential mobility models or the inherent incompleteness of real data. Microsimulation residential mobility models deal with complex behavioral processes and result in the development of particularly complex systems. During the development of such a system, model errors may be discovered, that is, errors at the theoretical, methodological or operational level of the model. Once model errors are corrected, all other errors are data errors, that is, errors that arise from the misspecification of the input data (parameters).

Both types of errors may be detected during the process of *calibration*. The behavior of the base population living in a geographical area over a time period, for which empirical data are available, is simulated. The outcome of the simulation is compared to the empirical data and the analysis of their variations reveals which code modules and parameters should be fine-tuned.

Unfortunately, no empirical data are available for the Municipality of Mytilene. Therefore, the process of calibration cannot be applied to the outcome of *URM-Microsim*. However, in order to show the function of the model namely, the interrelation between the development in demographic structures and the housing market, a series of simulations are carried out. Initially, a base run is executed using the population data created and the probability distribution calculated as described in **Chapter 6**. In the subsequent runs all probability distributions but one are held constant to show the effect of the changed one. In each subsequent run the values of *nest leaving*, *union formation* and *union dissolution* probabilities are doubled correspondingly. It should be noted that the simulations should be viewed in relation to each other. It is not the outcome of each simulation that is of importance, but rather the differences between the simulations and particularly between a given simulation and the base population run.

7.2.1 Base run simulation

The population database used in the simulation contains 29.085 individuals and 11.208 households. In Figure 58 and Figure 59 the map of Mytilene shows how these individuals and households are distributed in each residence zone correspondingly. In addition Figure 60 and Figure 61 show the number of dwellings and available dwellings of each residence zone correspondingly.

In Figure 62 and Figure 63 the map of Mytilene shows the number of in-migrating and out-migrating households per residence zone correspondingly. Moreover Figure 64 shows the number of households moving from each residence zone to Chrisomalousa (zone 1). Finally, in Figure 65 the map of Mytilene shows the range of households that entered minus the ones that left each residence zone of the city. In particular, the map shows that the recently developed, over-populated residence zones such as *Chrisomalousa* (zoneID = 1), *Kalithea* (zoneID = 2), *Gipedo* (zoneID = 3) and *Nosokomeio* (zoneID = 16) have a large number of incoming households as expected. One thing that is striking in the base run simulation results is that two old residence zones, *Kioski* (zoneID =22) and *Vounaraki* (zoneID =7), also have a large number of incoming households. This is due to the fact that a considerable number of old houses of these residence zones have been renovated and split to small one or two dwelling apartments (personal contact with civil engineers of Mytilene). In conjunction to fact that these residence zones have a low dwelling rent and market values, they attracted a large number of one and two member households (nest leavers and students) as shown in Table 39.

Table 40 gives an overview of the potential movers as well as the realized move in the base run. Union dissolution leavers and nest leavers are most successful in realizing the desired move since they have high search intensity and high offer acceptance probabilities. Out-migration is low almost in all event categories since households would prefer to substitute one or more of their search criteria to obtain a dwelling in the area of Mytilene than leave the city to search for a dwelling in another area. Those households that are slightly affected by demographic events, such as nest leaving stayers and union formation stayers have a low percentage of realized moves.

zoneID	Zone name	In_households – out_households	Household size
7	Vounaraki	**33**	1
7	Vounaraki	**29**	2
22	Kioski	**45**	1
22	Kioski	**20**	2

Table 39: Household change in 7 and 22 residence zones

	Event	event statistics	potential moves	internal realised moves	Out-migration	Total	% of realised moves
1	Death	281	199	138	12	**150**	75,38
2	fertility	442	389	231	51	**282**	72,49
3	nest leaving leaver	54	54	52	1	**53**	98,15
4	nest leaving stayers	54	47	12	3	**15**	31,91
5	union diss leaver	142	142	124	15[5]+3	**142**	100,00
6	union diss stayers	142	139	52	8	**60**	43,17
7	union form leavers	40	40	33	3	**36**	90,00
8	uinion form (one dwelling)	358	328	221	4	**225**	68,60
9	union form stayers	127	104	26	2	**28**	26,92
13	out-migration member	964	769	274	24	**298**	38,75
14	f-m form leavers	3	3	3	0	**3**	100,00
15	f-m form stayers	59	22	15	0	**15**	68,18
16	f-m form (one dwelling)	53	45	28	0	**28**	62,22
17	f-m leaving leaver	9	9	3	0	**3**	33,33
18	f-m leaving stayers	9	9	2	0	**2**	22,22
20	in-migration	700	700	700	0	**700**	100,00
	Total	**3437**	**2999**	**1914**	**126**	**2040**	
	number of households	12032					
	number of individuals	29103					
	number of in-migrating households	700					

Table 40: The number of potential movers and realized moves by event - base run

[5] There are 15 union dissolution leavers who out-migrate without first searching for a dwelling in the city. They are usually citizens of another Municipality.

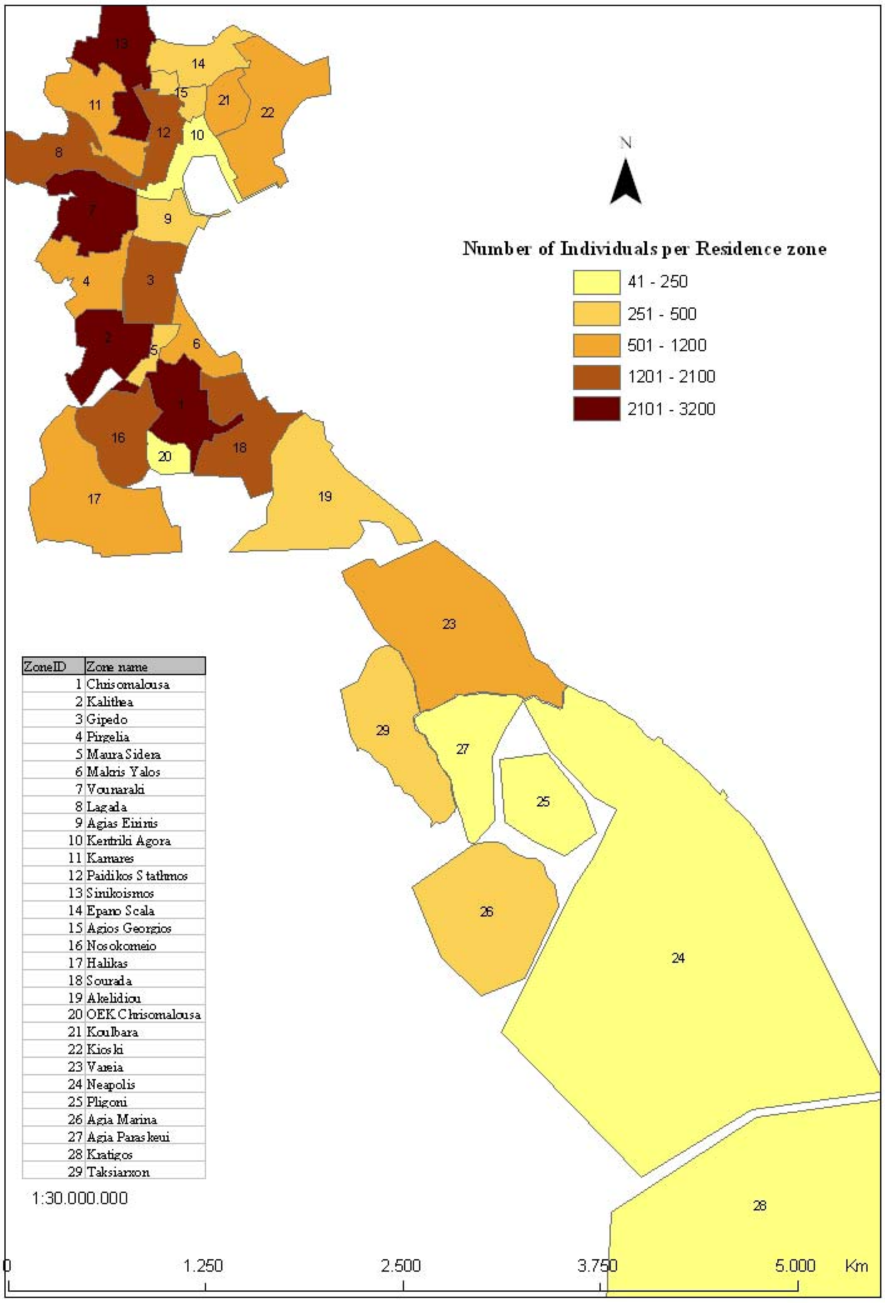


Figure 58: Number of Individuals per Residence zone – Map of Mytilene

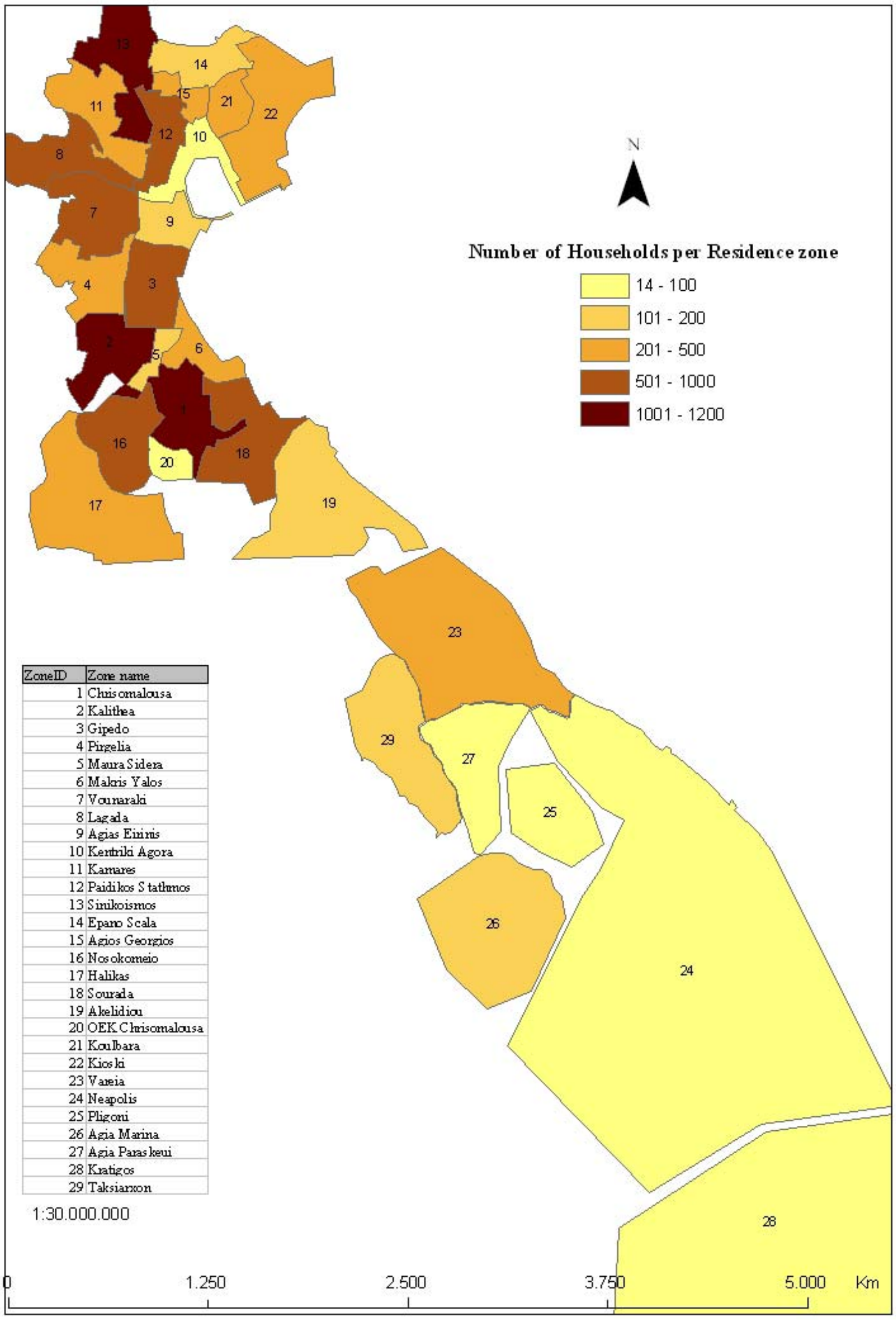


Figure 59: Number of Households per Residence zone – Map of Mytilene

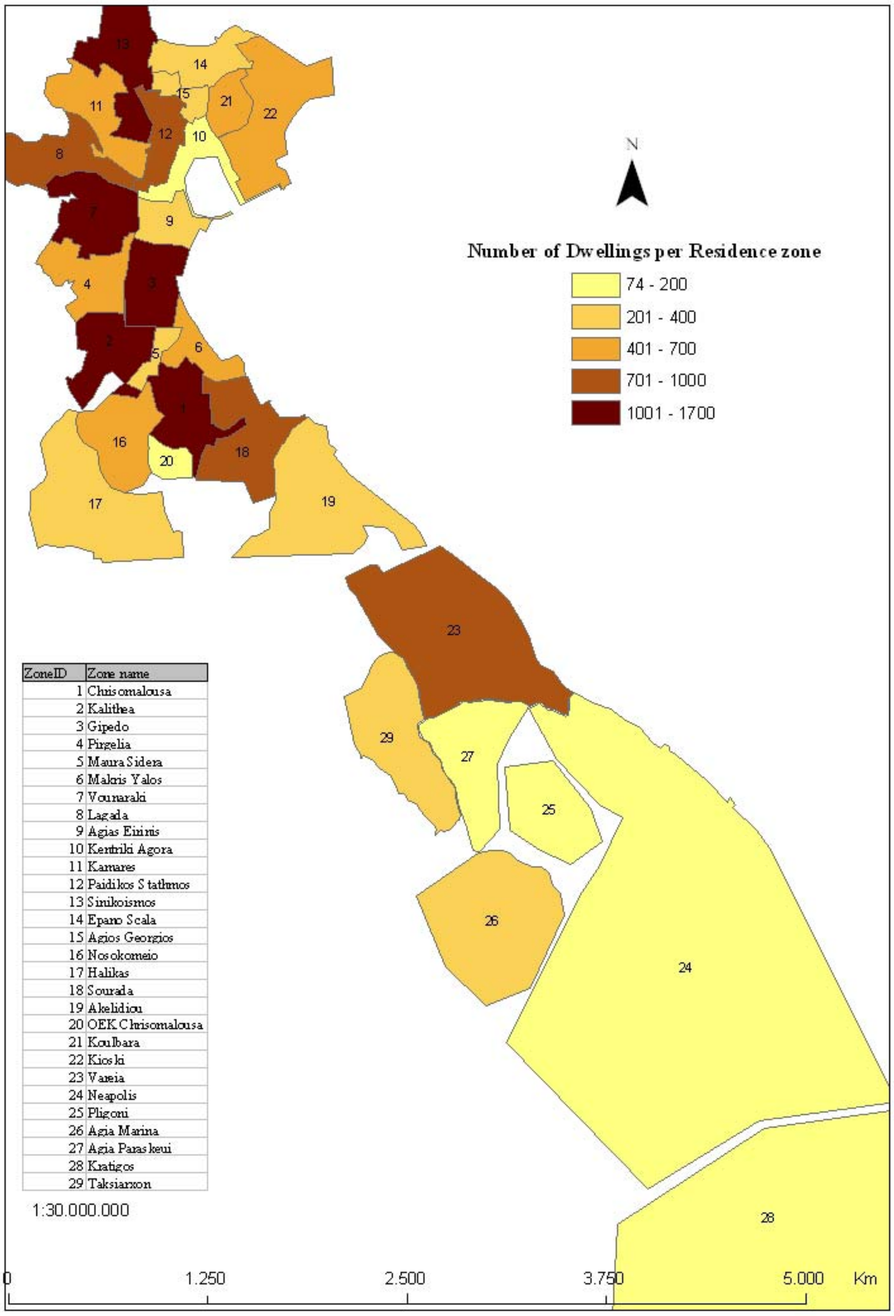


Figure 60: Number of Dwellings per Residence zone – Map of Mytilene

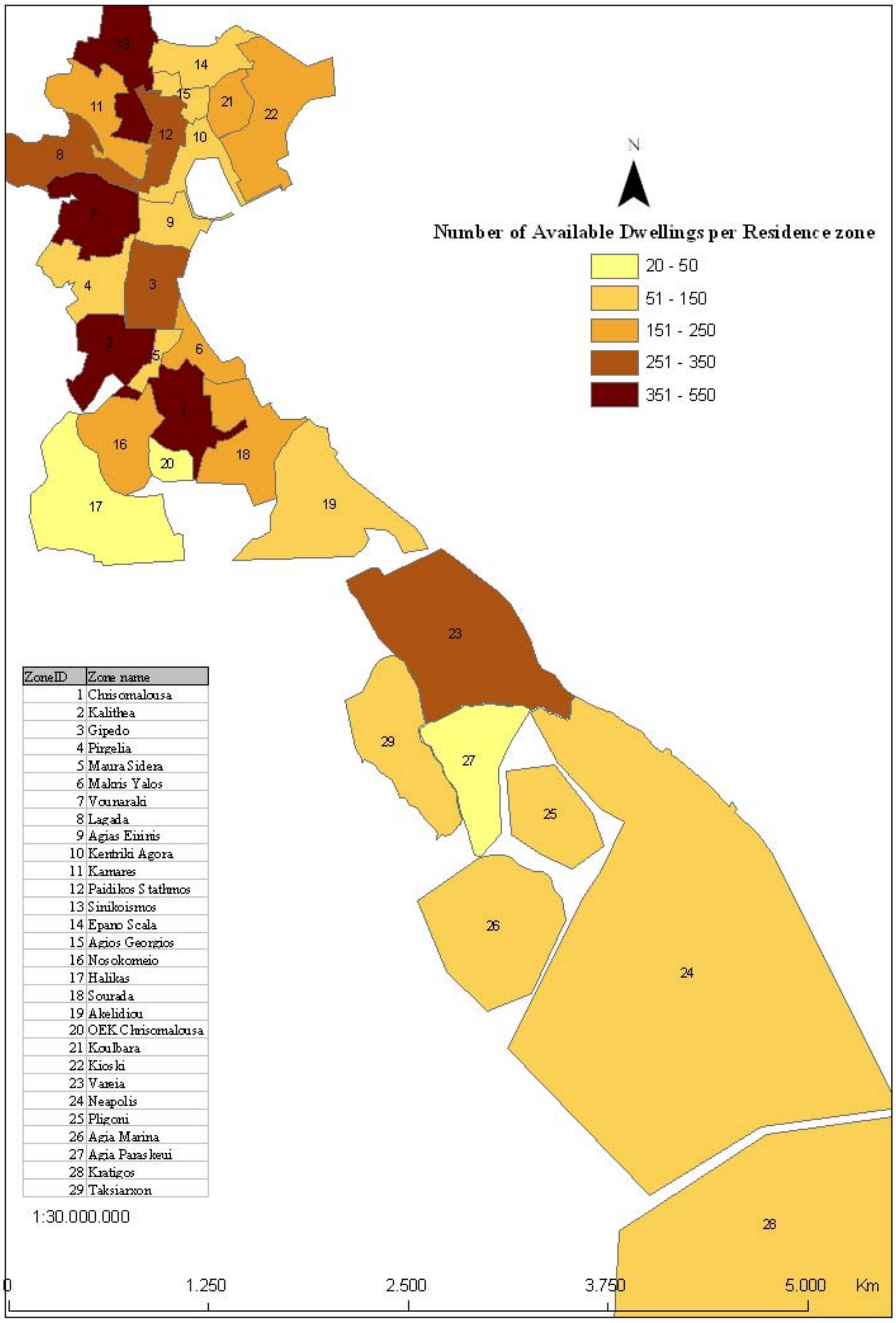


Figure 61: Number of Available Dwellings per Residence zone – Map of Mytilene

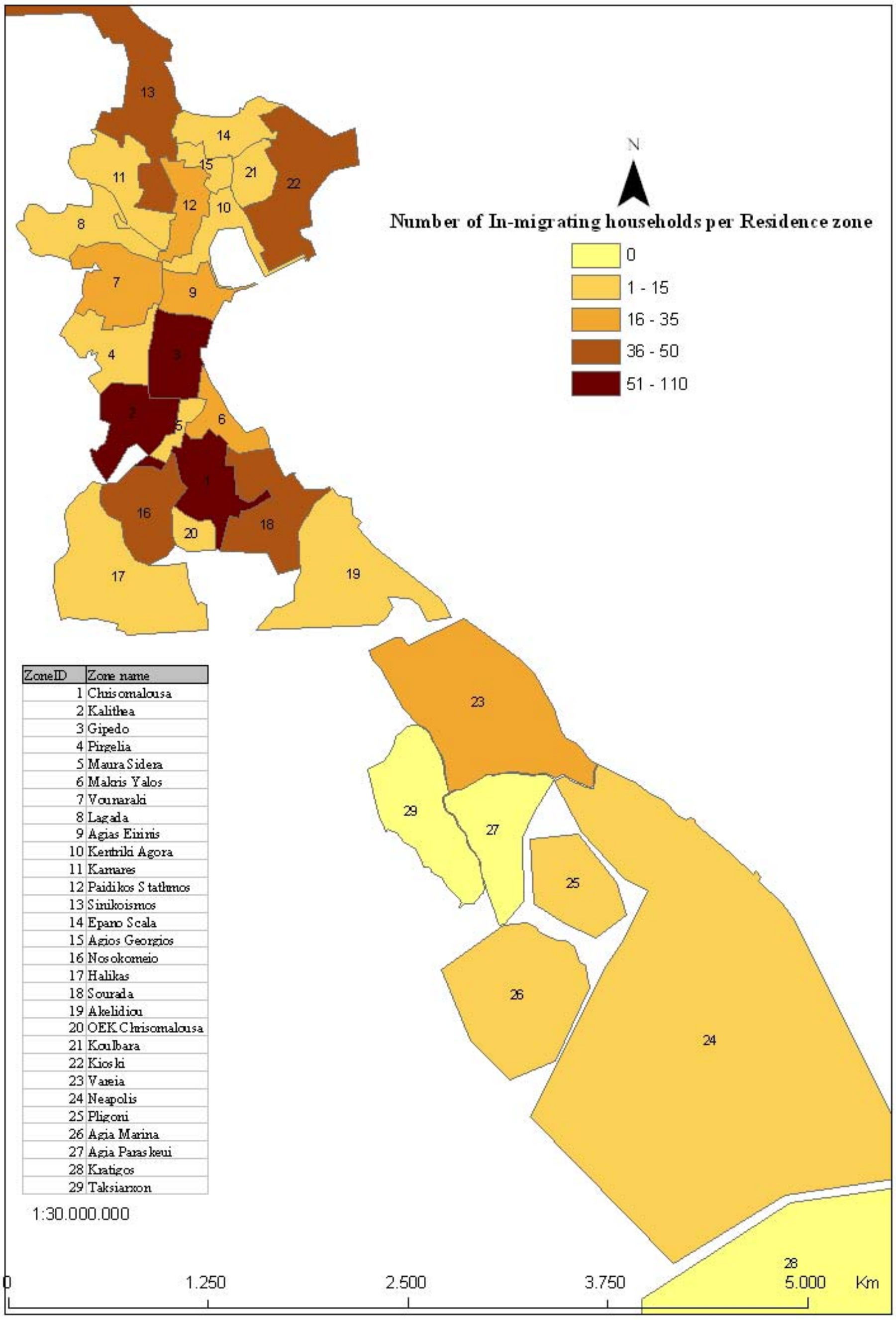


Figure 62: Number of In-migrating households per Residence zone – Map of Mytilene

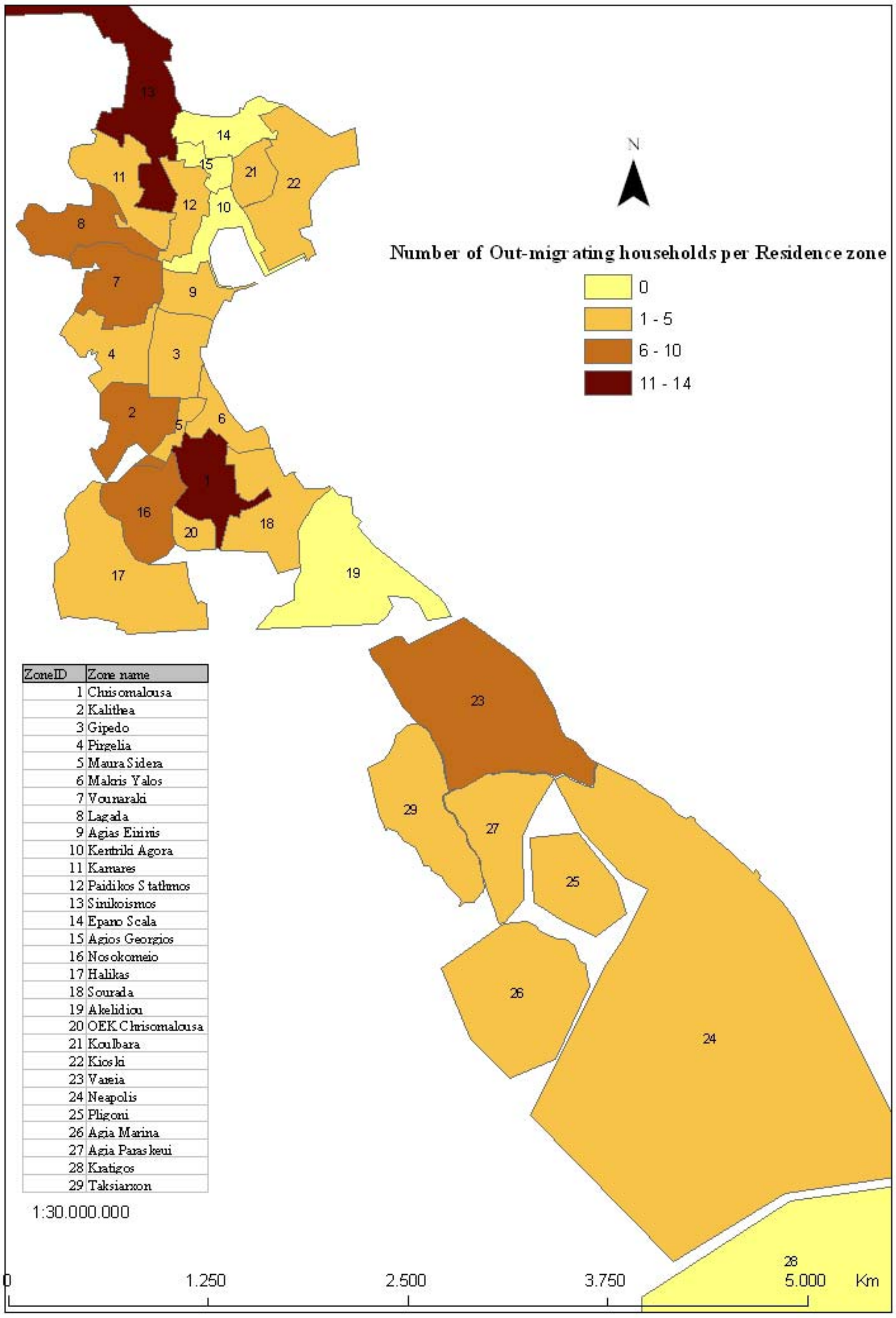


Figure 63: Number of Out-migrating households per Residence zone – Map of Mytilene

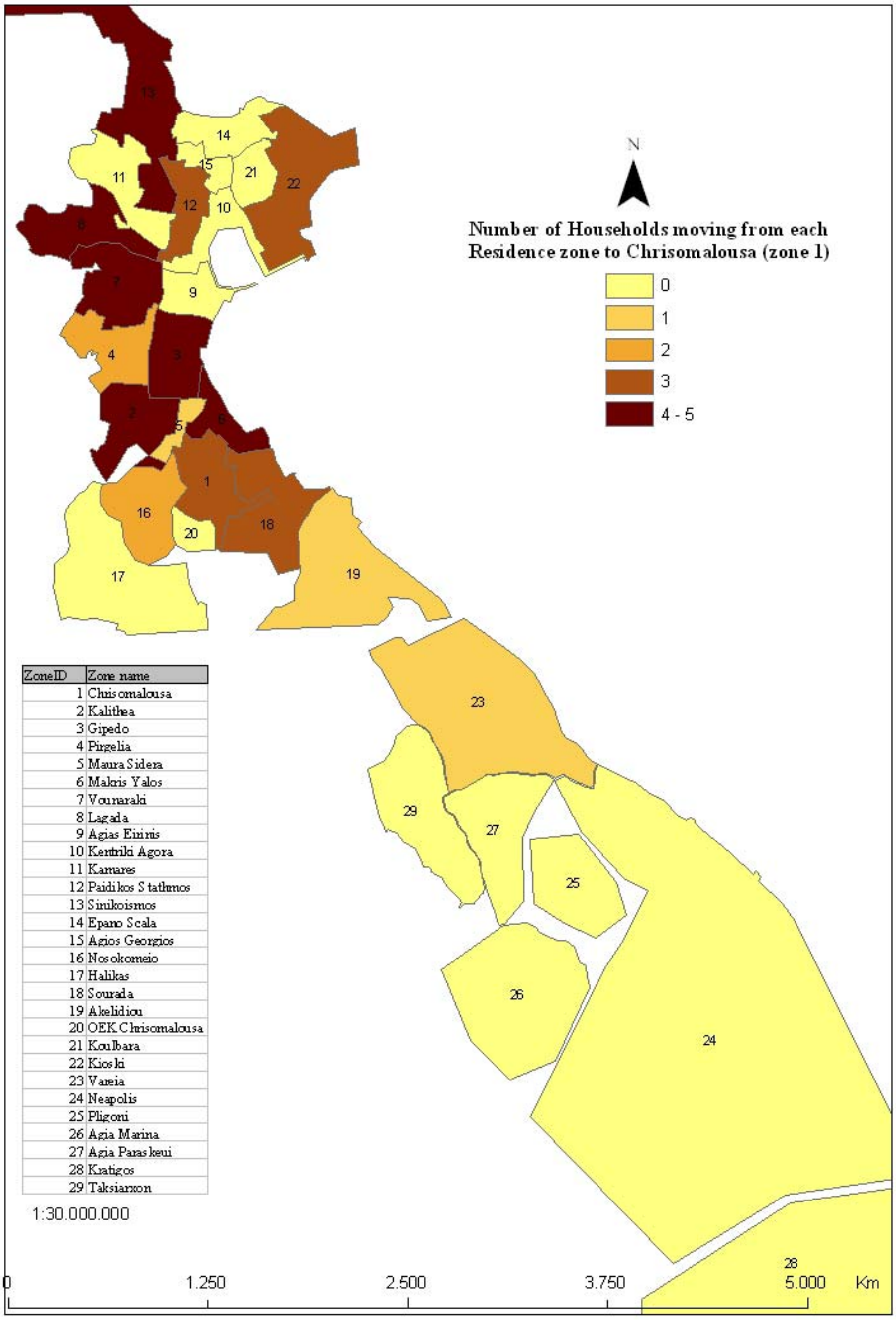


Figure 64: Number of Households moving from each Residence zone to Chrisomalousa (zone 1) – Map of Mytilene

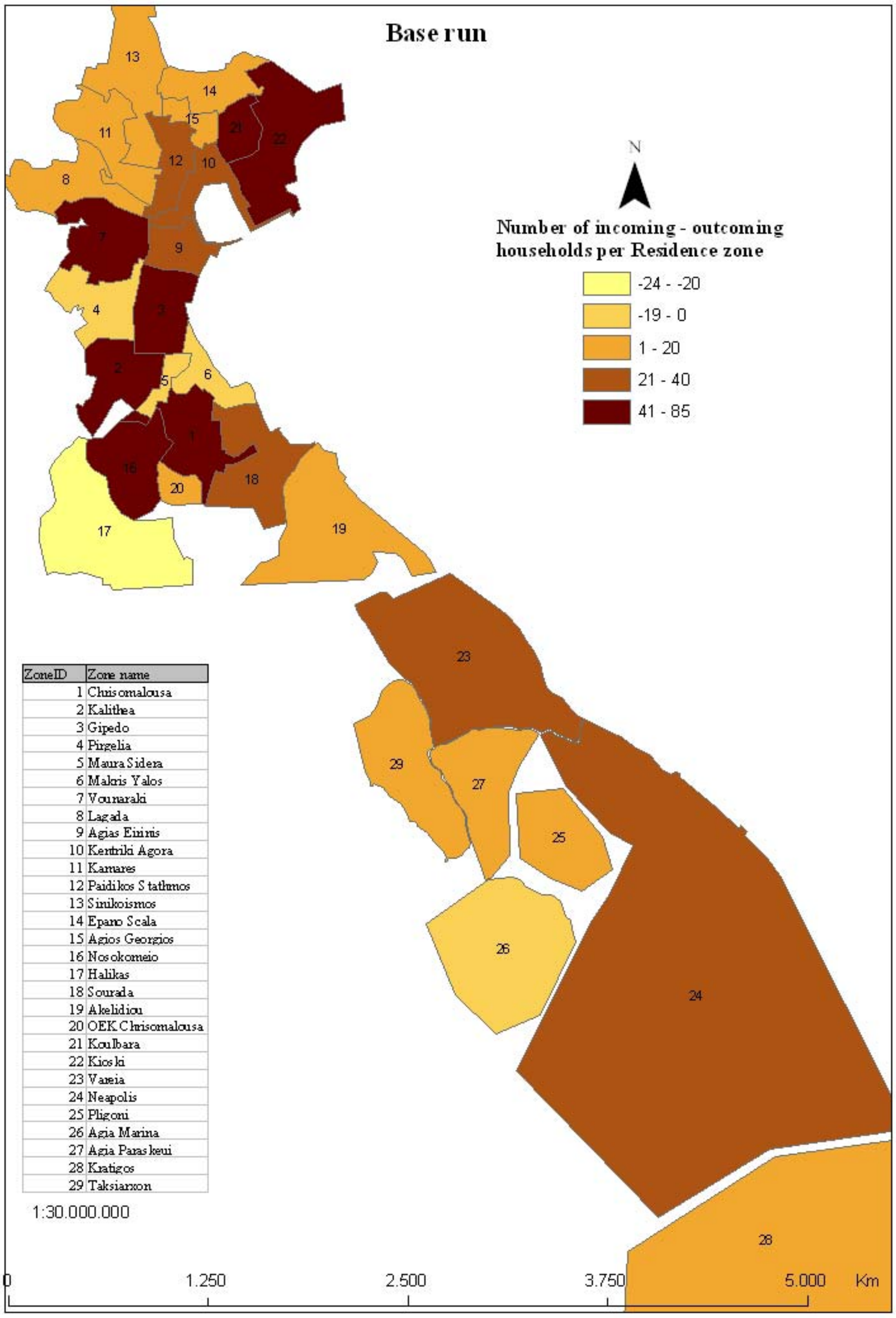


Figure 65: Base run simulation results – map of Mytilene

7.2.2 Nest leaving effect run simulation

To illustrate the effect of nest leaving, the nest leaving probability is doubled in the second simulation. As shown in Table 41 the number of nest leavers and consequently the number of nest leaving stayers are significantly increased: there are 116 potential nest leavers from 99 households instead of 54 out of 47 in the base run. In addition, the number of nest leavers' realized moves is also increased. In all, many more moves are made in this run than in the base run (2.136 instead of 2.040). This is due to the increased number of nest leavers who have high search intensity. In addition, the absolute number of realized moves of nest leavers is increased and consequently, more vacancies are created which gives moving households for reasons of death or other categories of stayers a better chance of success.

In Figure 66 the map of Mytilene shows the range of households that entered minus the ones that left each residence zone of the city. One may observe that the range of household moves of two residence zones with small, old houses and low dwelling rent/market values that is *Sinikoismos* (zoneID= 13) and *Lagada* (zoneID =8) has increased. This may be justified by the increase of nest leavers who prefer small and cheap dwellings as shown in Table 42.

	Event	event statistics	potential moves	internal realised moves	out-migration	Total	% of realised moves
1	death	287	214	161	18	**179**	83,64
2	fertility	421	368	212	43	**255**	69,29
3	**nest leaving leaver**	**116**	**116**	**111**	**5**	**116**	**100,00**
4	nest leaving stayers	116	99	27	4	**31**	31,31
5	union diss leaver	125	125	101	23+1	**125**	100,00
6	union diss stayers	125	119	45	7	**52**	43,70
7	union form leavers	42	42	36	3	**39**	92,86
8	uinion form (one dwelling)	370	341	229	1	**230**	67,45
9	union form stayers	134	105	30	1	**31**	29,52
13	out-migration member	1034	830	280	32	**312**	37,59
14	f-m form leavers	3	3	3	0	**3**	100,00
15	f-m form stayers	50	22	17	1	**18**	81,82
16	f-m form (one dwelling)	44	40	25	0	**25**	62,50
17	f-m leaving leaver	16	16	12	0	**12**	75,00
18	f-m leaving stayers	16	14	8	0	**8**	57,14
20	in-migration	700	700	700	0	**700**	100,00
	Total	**3599**	**3154**	**1997**	**139**	**2136**	
	number of households	12.094					
	number of individuals	28.999					
	number of in-migrating households	700					

Table 41: The number of potential movers and realized moves by event – effect of nest leaving

zoneID	Zone name	In_households – out_households		Household size
		Base run	Nest leaving run	
8	**Lagada**	**11**	**20**	1
13	**Sinikoismos**	**32**	**44**	1

Table 42: Base & nest leaving run household change in 8 and 13 residence zones

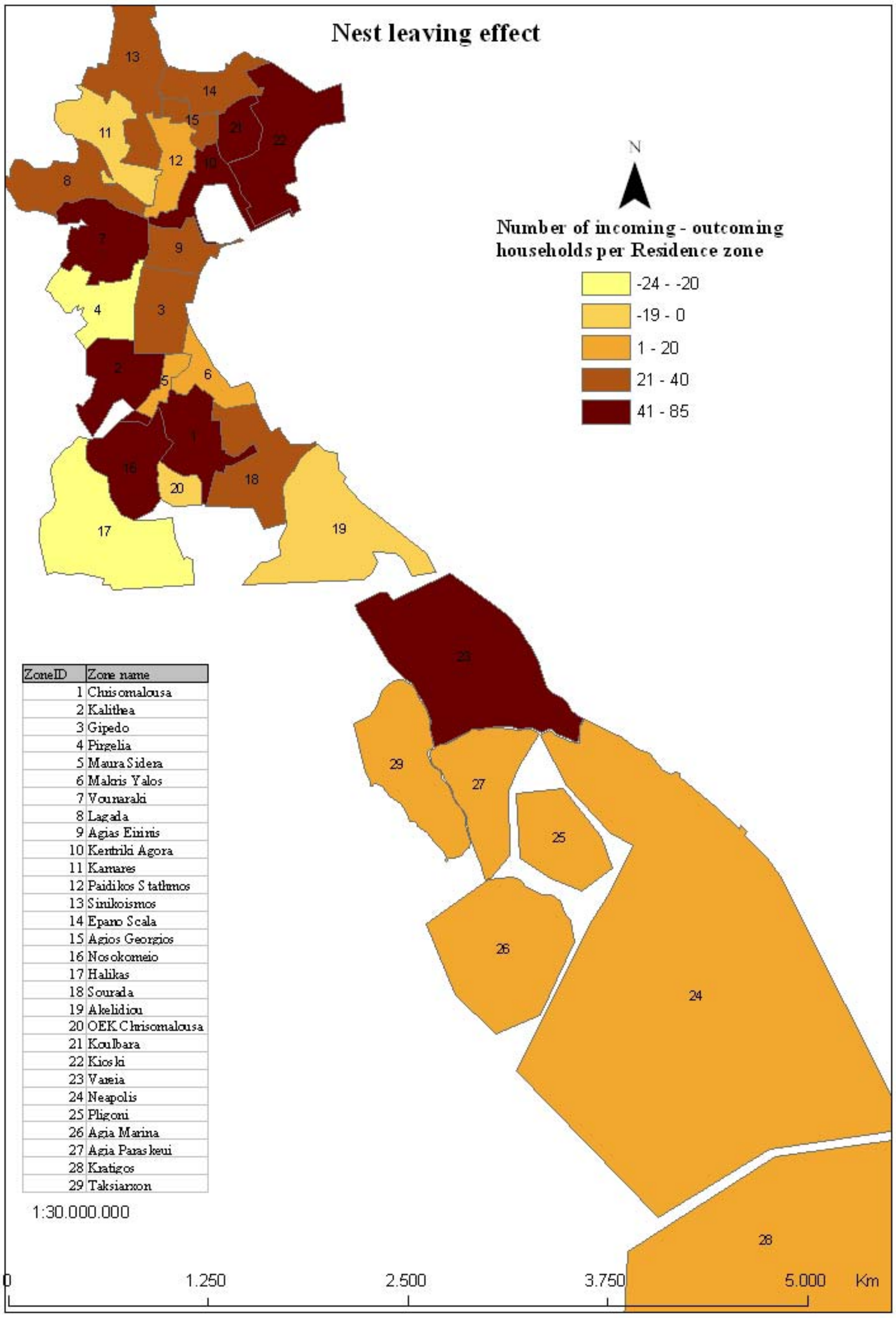


Figure 66: Nest leaving simulation results – map of Mytilene

7.2.3 Union dissolution effect run simulation

To illustrate the effect of union dissolution, the union dissolution probabilities are doubled in the third simulation. As shown in Table 43 the number of union dissolution leavers and consequently the number of union dissolution stayers are significantly increased: there are 250 potential union dissolution leavers from 327 households instead of 142 out of 139 in the base run. In addition, the number of union dissolution leavers' realized moves does not change significantly (from 100% to 99.60%). In all, many more moves are made in this run than in the base run (2.203 instead of 2.040). This is due to the increased number of union dissolution leavers who have high search intensity. In addition, the successful moves of nest leavers decrease from 98.15% in the base run to 94.74%. This is caused by the competition between nest leavers and union dissolution leavers in housing demand; both search for small houses (since in case of union dissolution the male usually moves out) and union dissolution leavers have higher search intensity than nest leavers.

In Figure 67 the map of Mytilene shows the range of households that entered minus the ones that left each residence zone of the city. One may observe that the divergence in the range of household moves of the residence zones smoothes out. For example, there are no residence zones with the lowest range of household moves (-29 to -17). This may be justified by the increase of union dissolution leavers who come from all social and economic layers and thus, although they prefer small dwellings they may afford to live in high rent/market residence zones as shown in Table 44.

	Event	event statistics	potential moves	internal realised moves	Out-migration	Total	% of realised moves
1	Death	294	229	161	14	**175**	76,42
2	Fertility	412	357	220	34	**254**	71,15
3	nest leaving leaver	57	57	54	0	**54**	94,74
4	nest leaving stayers	57	53	23	1	**24**	45,28
5	**union diss leaver**	**250**	**250**	**211**	**37+1**	**249**	**99,60**
6	union diss stayers	250	237	90	9	**99**	41,77
7	union form leavers	37	37	29	5	**34**	91,89
8	uinion form (one dwelling)	327	307	206	6	**212**	69,06
9	union form stayers	121	100	33	3	**36**	36,00
13	out-migration member	1036	805	276	35	**311**	38,63
14	f-m form leavers	6	6	6	0	**6**	100,00
15	f-m form stayers	70	32	22	0	**22**	68,75
16	f-m form (one dwelling)	58	44	20	0	**20**	45,45
17	f-m leaving leaver	11	11	5	0	**5**	45,45
18	f-m leaving stayers	11	9	2	0	**2**	22,22
20	in-migration	700	700	700	0	**700**	100,00
	Total	**3697**	**3234**	**2058**	**145**	**2203**	
	number of households	12.136					
	number of individuals	28.961					
	number of in-migrating households	700					

Table 43: The number of potential movers and realized moves by event – effect of union dissolution

Number of Union dissolution leavers	Number of rooms	Dwelling rent value	Dwelling market value
55	1	260	18200
30	1	280	24000
6	1	300	29000
2	1	350	24000
82	1	450	30000
12	2	280	36400
3	2	450	60000
1	3	300	54600
1	3	400	87000
2	3	450	90000
4	4	280	96000
1	4	350	72800
5	4	450	120000
1	5	370	91000
4	5	450	150000
1	7	450	210000

Table 44: Number of rooms and rent/market value of dwellings chosen by union dissolution leavers

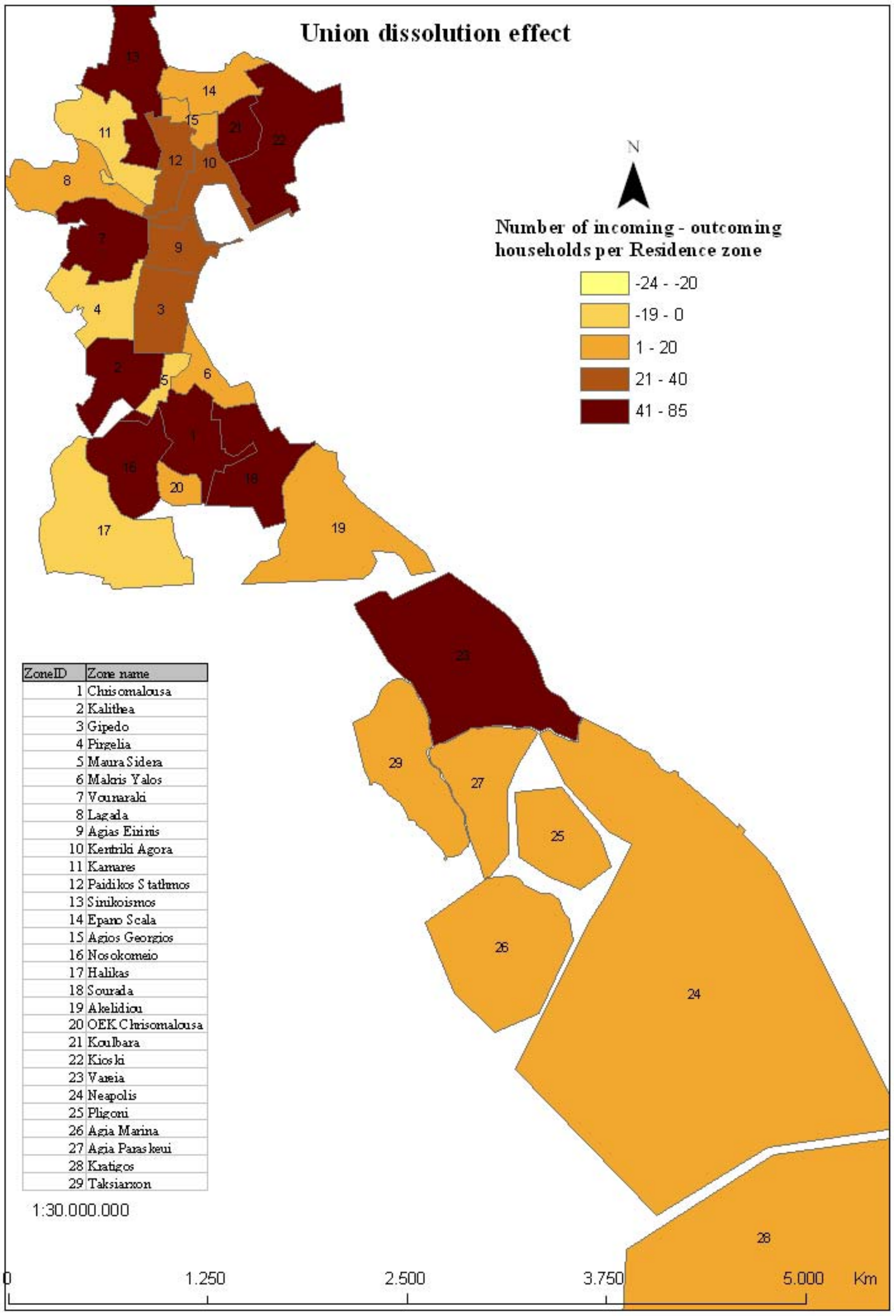


Figure 67: Union dissolution simulation results – map of Mytilene

7.2.4 Union formation effect run simulation

To illustrate the effect of union formation, the union formation probabilities are doubled in the forth simulation. As shown in Table 45 the number of union formation leavers as well as union formation (one dwelling) and consequently the number of union dissolution stayers are significantly increased: there are 128 potential union formation leavers and 387 union formations with at least one dwelling from 312 households instead of 40 and 328 out of 104 in the base run. Union formations with at least one dwelling are not significantly increased since cohabitants who get married are also included in this type of event and they are the same in number in both runs. In addition, the number of union formation leavers' realized moves slightly decreases (from 90% to 86.72%). In all, many more moves are made in this run than in the base run (2.189 instead of 2.040). This is due to the increased number of union formation leavers who have high search intensity. In addition, the successful moves of nest leavers decrease from 98.15% in the base run to 87.27%. This indicates that there is competition between nest leavers and union formation leavers in housing demand; both search for small houses.

In Figure 68 the map of Mytilene shows the range of households that entered minus the ones that left each residence zone of the city. One may observe that there is a slight increase in the range of household moves of residence zones with high rent/market values such as *Sourada* (zoneID =18) and *Vareia* (zoneID =23). This may be justified by the increase of union formation leavers who although come from all social and economic layers, they prefer newly constructed and consequently, more expensive dwellings as shown in Table 46. It should be noted that most newly constructed dwellings reside in the recently developed and consequently, more expensive residence zones.

	Event	event statistics	potential moves	internal realised moves	out-migration	Total	% of realised moves
1	Death	287	211	144	15	**159**	75,36
2	Fertility	401	347	213	34	**247**	71,18
3	nest leaving leaver	55	55	47	1	**48**	87,27
4	nest leaving stayers	55	48	21	0	**21**	43,75
5	union diss leaver	141	141	119	20+2	**141**	100,00
6	union diss stayers	141	135	45	5	**50**	37,04
7	**union form leavers**	**128**	**128**	**109**	**2**	**111**	**86,72**
8	uinion form (one dwelling)	415	387	254	12	**266**	68,73
9	union form stayers	367	312	101	6	**107**	34,29
13	out-migration member	975	761	251	26	**277**	36,40
14	f-m form leavers	3	3	3	0	**3**	100,00
15	f-m form stayers	61	24	15	0	**15**	62,50
16	f-m form (one dwelling)	55	47	34	0	**34**	72,34
17	f-m leaving leaver	12	12	8	0	**8**	66,67
18	f-m leaving stayers	12	9	1	1	**2**	22,22
20	in-migration	700	700	700	0	**700**	100,00
	Total	**3808**	**141**	**2065**	**124**	**2189**	
	number of households	12.115					
	number of individuals	29.043					
	number of in-migrating households	700					

Table 45: The number of potential movers and realized moves by event – effect of union formation

Number of Union formation leavers	Number of rooms	Dwelling rent value	Dwelling market value
3	1	280	24000
1	1	300	29000
10	1	450	30000
24	2	280	36400
1	2	300	48000
5	2	350	58000
45	2	450	60000
2	3	400	87000
5	3	450	90000
1	4	370	96000
1	4	420	116000
5	4	450	120000
6	5	450	150000

Table 46: Number of rooms and rent/market value of dwellings chosen by union formation leavers

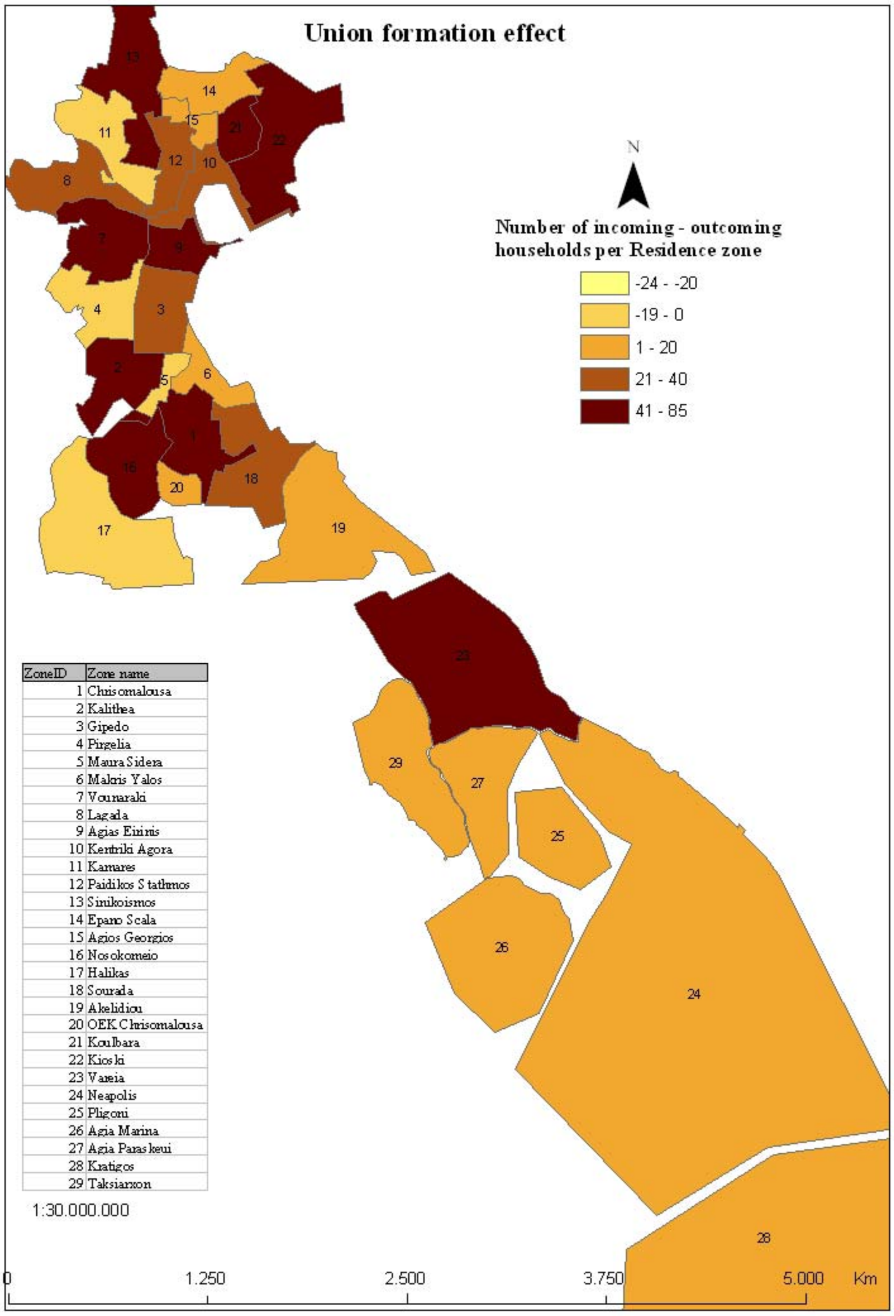


Figure 68: Union formation simulation results – map of Mytilene

7.3 Validation

Once *URM-Microsim* is calibrated it should be validated to verify that the model may be successfully applied in another context. In other words, the process of *validation* applies the calibrated model to different data and/or geographical area and/or period, and checks whether it yields comparable results i.e. whether the model 'fits' in another context (Oskamp, 1997).

According to (Oskamp, 1997), four types of validation may be discerned. In in-sample validation, simulation results on a given sample are tested against the empirical values of the same sample. Conventional test-statistics for measures of goodness of fit are calculated (R^2, t- and F-statistics). In out-of-sample validation, the model is tested against a different sample of the base population aiming to show whether the model successfully simulates various samples. The same statistical measures as in the in-sample validation are used. In out-of-type validation the models is tested against a different data source. Finally, in out-of context validation, different data from a different geographical, temporal or institutional context are entered to verify that the model also performs in another context.

Unfortunately, no empirical data are available for the Municipality of Mytilene. Therefore, none of the first three types of validation can be applied to the outcome of *URM-Microsim*. Since the validation of the model is of great importance, the use of out-of-context validation, i.e. the loading of *URM-Microsim* with data from another city, is part of this thesis' future work.

7.4 Summary

This chapter focused on the performance of *URM-Microsim*. Once the system is loaded with the Mytilene data, it is executed in order to simulate the population and housing market developments. Then, the issue of model efficiency arises: how well does the model simulate reality? Initially, the notion of system calibration was discussed then the series of simulations carried out for *URM-Microsim* were presented and the simulation results were discussed. Next, the notion and methods of system validation were discussed.

Chapter 8
Conclusions and Future Research

8.1 Introduction

This chapter provides an overview of the work presented in this thesis, it summarises the main contributions of the research and presents its conclusions. In addition, further work in this area is discussed.

8.2 Thesis overview

URM-Microsim (Urban Residential Mobility Microsimulation) is a dynamic, spatio-temporal microsimulation prototype decision-support system that captures the evolution of the housing market of an urban area on a yearly basis, taking account of housing demand and supply. The system is capable of evaluating the impact of demographic and economic scenaria on the building area of a city.

Within *URM-Microsim*, the database entities of interest are *individuals*, *households* and *dwellings*. The population of a city consists of *households*. A household has one or more members. Each household member is an *individual* with a specific position in the household (couple member, child or single). An individual may have an *occupation* (and therefore, a salary) or not. The occupation belongs to a *residence zone*. A household lives at a *dwelling* that owns or rents. A dwelling belongs to a *building*, which in turn belongs to a *building block*, which in turn, belongs to a *residence zone*. In addition, the database entities have characteristics, whose values may change over time as well as the relationships between the entities.

The model consists of three sub-models (*packages*): *Housing Demand*, *Housing Supply*, and *Residential Search and Migration*. The *Housing demand* sub-model simulates the occurrence of demographic and economic events (for example death, fertility, marriage, occupation and income change) and therefore determines the number of households that will search for a new house. The *housing supply* sub-model on the other hand, simulates the changes in housing stock and therefore determines the number of houses that are available to households. Finally, the *Residential Search and Migration* sub-model simulates the clearing of the housing market, that is, which households actually find a new dwelling. The outcome of the *Residential Search and Migration* sub-model affects the way housing demand and supply are modified. Three types of move by destination are distinguished: households moving within the boundaries of the housing market (*residential mobility*), households moving from the housing market to another one (*out-migration*), and households moving into the housing market from the rest of the world (*in-migration*). Most of the processes are simulated using *Monte Carlo techniques*. In some cases, the process is clearly one of choice and may therefore be modeled with a decision rules approach. The order of events (sequencing) occurring during *a one-year period* is simulated using *random dates*. The probability distributions of the various events can be age, sex or duration specific.

The *Housing Demand* sub-model incorporates three types of events: *Demographic change*, *Income change* and *Demand change* events.

- *Demographic change events.* Each individual is subjected to the probability of experiencing one of the *death*, *fertility, union formation, union dissolution, flat-mate formation, flat-mate leaving, nest leaving* events and *out-migration*, based on the individual's age and stage in life – cycle. In addition, the number of *in-migrating* households is determined. Finally, an individual or a household may decide to *out-migrate* due to several reasons. For example, an individual who is unemployed for a long time may decide to out-migrate.

- *Income change events*. The *education* group to which the individual belongs is determined. Next, each individual of age 15 and above, who does not work (soldiers are excluded) is subjected to the probability of *getting a job* based on the individual's sex, age, current employment. If the individual already works he/she is subjected to the probability of *job changing*, retiring or becoming unemployed based on the individual's sex, age, current employment and education status. Finally, each individual of age 18 and above is subjected to the probability of having his/her *income changed* based on the individual's education career and occupational changes.
- *Demand change event.* A household that does not have experienced a demographic or income change event is subjected to the probability of demand changing.

The *Housing supply* sub-model incorporates four types of events: *New construction*, *Structure conversion*, *Demolition* and *Housing expenditures change*.

- *New construction event.* The number of new dwellings to be added to the house stock is determined based on the market's mobility and housing shortage. The location of each new dwelling is determined based on the undeveloped land that is suitable for development, the market area and the market's housing demand trends. The dwelling's type and services are determined based on the market's housing demand trends.
- *Structure conversion events.* The possible structure changes of a dwelling are *split*, *combination* or *renovation*. Each dwelling is

subjected to the probability of being split, combined or renovated, based on the cost and difficulty of the each conversion.

- *Demolition event.* A dwelling in poor condition is subjected to the probability of being demolished. If the dwelling is occupied, then the corresponding household must search for another dwelling. The demolition takes place only after the household moves out.
- *Housing expenditures change event.* The housing expenditures (maintenance capital, heat, electricity etc) of each dwelling are determined based on the dwelling's type, size, age and construction quality.

Households that are dissatisfied with their current housing, *search* the *housing market* for a vacant dwelling that meets their needs. If a household finds a dwelling that complies with its preferences, then the household *moves* into the new dwelling. If household search fails, then the household either *remains* to its current dwelling or migrates to another housing market (*out-migration*). Housing demand and supply determine the rent and market value of the dwellings (*housing prices*). The sub-model incorporates the *Search and Migration* and *Housing prices* events.

- *Search and migration event.* All households whose structure or housing circumstances have changed, all newly formed households, as well as the households they originate from are simulated to search for a new dwelling. The household searches the list of vacant dwellings for a dwelling according to the following rules:
  - *Tenure*. Initially, the household decides whether it will search for a house to own or rent, based on the household's *previous tenure*, its current *income*, and its current *life cycle stage.*
  - Dwellings are defined by seven (7) dimensions: *dwelling price/rent, number of rooms, dwelling type, dwelling age*, *structure condition,*

amenities (garden, parking etc), and *neighborhood quality.* The household defines the values of the dimensions the desired dwelling should have, as well as the dimensions that are of importance (primary dimensions) and those that are not (secondary dimensions). The desired dwelling is defined based on the event experienced by the household, the household's income as well as whether the household searches for a dwelling to rent or own.

  - The *number of dwellings considered* by the household depends on the *search intensity* of the household. In other words, households with high search intensity consider more dwellings than the ones with low search intensity.
  - If the household decides to accept a vacancy then it moves to the new dwelling and the dwelling the household occupied (if any) is vacated (*Dwelling vacation).* If the household fails to find a dwelling, then if it is an in-migrating household, the household as well as its members is deleted. In any other case, the household is subjected to the probability of *out-migrating*. If the household does not out-migrate then if it already has a dwelling it remains there, or in any other case the household is deleted and its members return to their parental households.

- *Housing prices change event.* Once all household moves have taken place, the rent and market values of the dwellings are updated based on housing demand and supply.

URM-Microsim was implemented and tested within the city of Mytilene, using existing data.

8.3 Thesis limitations

The need to mitigate the repercussions of urban sprawl over the last decades triggered the use of systematic urban plans that encompass land use, transportation and environmental dimensions. These plans aim to forecast housing demand and to analyze the impacts governmental policy programs may have on urban development. While a considerable number of transportation models has been developed and used in the regional transportation planning process, research review revealed that land use models have received little attention. The lack of a fully documented and easily expandable land use model that incorporates spatial representation of model results at various levels of aggregation triggered the need for the development of *URM-Microsim*.

URM-Microsim fulfils all thesis objectives discussed in section 1.2. It incorporates most of the elements of a residential mobility microsimulation model discussed in section 2.5 (it has not been fully calibrated and validated yet). Therefore, it is theoretically sound and provides the capacity to analyze the impact of the events experienced by individuals and households in population residential mobility and consequently in the change of a city structure.

In addition, its development followed the STUML process, a set of activities that leads to the orderly construction of the system under development by applying the UML modeling techniques extended with spatiotemporal constructs. The use of a process during all phases of system development in conjunction with the use a graphical modeling language during analysis and design make it possible for a researcher to comprehend *URM-Microsim* development even if he/she does not have a programmer's background. That makes *URM-Microsim* easily extendable with additional functionality and easily adaptable to the cultural, social, economic and geographical peculiarities of each city.

Moreover, the use of the ArcGIS platform for the implementation of *URM-Microsim* allows the spatial representation of model results at various levels of aggregation i.e. at building block and residential zone levels.

Finally, the use of the developed system was demonstrated by applying it to the city of Mytilene. For this purpose *demographic and housing market data* for the city of Mytilene was gathered from various data sources (e.g. National Statistical Service of Greece, Eurostat, Surveys etc) and was inserted in the developed system, in order to simulate the residential mobility of Mytilene. Unfortunately, a full calibration of *URM-Microsim* was not possible since no empirical data were available for the city of Mytilene.

However, it should be noted that during the course of model development, parameter estimation and model application, a number of *URM-Microsim* elements were under- or undeveloped. First, the Housing Supply sub-model was not implemented. All new dwellings and structure conversions are given exogenously. In addition, rent and market value changes of the dwellings are also given exogenously. Second, the income event was not implemented and households do not search for a new dwelling due to income changes. Third, in job entry and job change events the location of employment is not taken into consideration. Finally, in-migration event takes into consideration only one type of in-migrating households: students.

8.4 Future work

URM-Microsim is a prototype system and thus is susceptible of a considerable number of improvements and extensions.

First, *URM-Microsim* should be fully calibrated and then validated. The importance of calibrating and validating complex systems such as a residential mobility system was outlined in **Chapter 7**. Therefore, *URM-Microsim* should be applied to different data from a different geographical, temporal or institutional context are entered to verify that the model also performs in another context.

Second, some of the events modelled in *URM-Microsim* may be improved by further elaborating their design, implementation or parameter estimation: partner search, housing supply, income and in-migration.

In *URM-Microsim* an individual who is not married searches for a partner based only on age since only this information was available. Evidently, the search criteria for union formation are more, such as education level, social and economical status and nationality.

The lack of digitized information concerning housing construction in Greece as well as the lack of organized research data concerning urban development inhibited the implementation of the Housing Supply sub-model. All new dwellings and structure conversions are given exogenously. In addition, rent and market values changes of the dwellings are also given exogenously. Since housing supply develops on the long-term to satisfy housing demand, and rent and market values of the dwellings depend on housing demand and supply, they should be incorporated in *URM-Microsim.*

The analysis and design models of *URM-Microsim* draw on the notion that income changes may trigger mobility. In order to implement this notion income changes at the individual level, or at least at disaggregate level are required. Since this information was not available, the income event was not implemented. Therefore households do not search for a new dwelling due to income to changes.

The analysis model of *URM-Microsim* draws on the notion that an individual searches for a job according to the certain criteria: However, since data statistics are not available, this notion was dropped during design and the individual is subjected to the probability of *job entering* or *changing*, based on the his/her sex, age, current employment (if any) and education status. In addition, the location of employment is not taken into consideration due to lack of information concerning job vacancies and work places.

In the implementation of *URM-Microsim*, *in-migration* event takes into consideration only one type of in-migrating households: students. This is due to the lack of any other statistics available. Other categories of in-migrating households should be taken into consideration.

Finally, the dwelling search algorithm should be elaborated to incorporate all search dimensions. Currently the most important search dimensions are considered (primary dimensions), namely *size (number of rooms)*, *price*, *type* and *tenure*. No statistics were available to determine the remaining three dimensions of a dwelling (secondary dimensions): *structure condition, amenities (garden, parking etc),* and *neighborhood quality*.

Third, a number of parameters were found to be inadequate or even non-existent. For example, since no statistical data are available, the probabilities of out-migration of unsuccessful searchers based on the event experienced are assumed, taking into consideration the fact that households would substitute one or more of their search criteria to obtain a dwelling in the area of Mytilene. **Chapter 6** describes all calculations and assumptions made for probability distributions.

8.5 Epilogue

In this concluding chapter, an overview of *URM-Microsim* was given and possible enhancements and directions for future work were also discussed.

The innovation of the *URM-Microsim* lies in the use of the most effective data modeling technique i.e. microsimulation, in conjunction with the most effective system development tools i.e. spatio-temporal UML and GIS. The outcome is a model that handles the inherent complexity of residential mobility systems effectively and at the same time is transparent and consequently comprehendible to all actors involved. This makes *URM-Microsim* easily expandable.

Model development of a complex system such as residential mobility systems is never completed. Improvements, extensions and updates can be made continuously. Future work on *URM-Microsim* should focus on broadening the scope of the model, adding and improving modules and parameters, and elaborating on the analysis and design models underlying *URM-Microsim.*

Bibliography

Abu-Lughod, J.L. and M.M. Foley (1960). Consumer strategies. Housing choices and housing constraints, Part 2, 387-447. N.N. Foote, J.L. Abu-Lughod, M.M. Foley and L. Winnick, editors. New York: McGraw-Hill.

Bartlema, J., Vossen, A. (1988). Reflections on household modelling. In Modelling household formation and dissolution, 243-253. N. Keilman, A. Kuijsten and A. Vossen, editors. Oxford: Clarendon.

Bedard, Y. (1999). Visual Modeling of Spatial Databases: Towards Spatial PVL and UML. Geomatica. Vol. 53, No.2, pp. 169-185.

Bedard, Y., Proulx, M.J. (2000). Perceptory. Centre de recherche en geomatique, Universite Laval, Quebec, Canada. (http://sirs.scg.ulaval.ca/Perceptory).

Birkin, M., Clarke, M., Chen, H., Dew, P., Keen, J., Rees, P., Xu, J. (2005). MOSES: Modeling and Simulation for e-Social Science. Proceedings of the AHM.

Birkin, M., Turner, A., Wu, B. (2006). A Synthetic Demographic Model of the UK Population: Methods, Progress and Problems. Proceedings of the 2nd International Conference on e-Social Science, Manchester, 28 - 30 June.

Booch, G. (1994). Object-Oriented Analysis and Design with Applications. Benjamin/ Cummings Publishing Company.

Booch, G., Rumbaugh, J., Jacobson, I. (1999). The Unified Modeling Language User Guide. Addison-Wesley Longman Inc.

Βουνάτσου, Β. (2002). Κτηματαγορά και η ανάπτυξη μιας μεσαίου μεγέθους πόλης: η περίπτωση της Μυτιλήνης. Πτυχιακή εργασία. Τμήμα Γεωγραφίας, Πανεπιστήμιο Αιγαίου.

Brodeur, J., Bedard, Y., Proulx M.J. (2000). Modeling Geospatial Application Databases using UML-based Repositories Aligned with International Standards in Geomatics. Geomatics, ACMGIS 2000. Washington DC, USA, 10-11 November.

Brown, H.J. (1975). Changes in workplace and residential locations. Journal of the American Institute of Planners 41, 32-39.

Brown, L.A., Moore, E.G. (1970). The intra-urban migration process: a perspective. Geografiska Annaler B 52, 1-13.

Burgess, E.W. (1925). The growth of the city: an introduction to a research project. The city. 47-62. R.E. Park, E.W. Burgess and R.D. McKenzie, editors. Chicago: University of Chicago Press.

Clarke, M. (1995). A Microsimulation approach to demographic and social accounting, in Social and demographic accounting, 195-221. G.J.D. Hewings and M. Madden, editors. Cambridge: Cambridge University Press.

Clarke, M., Longley, P.A., Williams, H.C.W.L. (1989). Microanalysis and simulation of housing careers: subsidy and accumulation in the UK housing market. Papers of the Regional Science Association 66, 105-122.

Clarke, M., Holm, E. (1987). Microsimulation methods in spatial analysis and planning. Geografiska Annaler B 69, 145-164.

Clark, W.A.V., Deurloo M.C., Dieleman F.M. (1986). Residential mobility in Dutch housing markets. Environment and Planning A 18, 763-788.

Clark, W.A.V. (1982). Recent research on migration and mobility: a review and interpretation. Progress in Planning 18, 1-56.

Connerly, C.E. (1986). The impact of neighbourhood social relations on prospective mobility. Social Science Quarterly 67, 186-194.

Devisch, O.T.J., Timmermans, H.J.P., Arentze, T.A., Borgers, A.W.J. (2004). Towards a Generic Multi-Agent Engine for the Simulation of Spatial Behavioral Processes. In: Van Leeuwen, J.P. and H.J.P Timmermans (eds.) Recent advance in Design & Decision Support Systems in Architecture and Urban Planning. Dordrecht: Kluwer Academic Publishers, ISBN: 1-4020-2408-8, p. 145-160.

Duncan, G.J. and S.J. Newman (1976). Expected and actual residential mobility. Journal of the American Institute of Planners 42, 174-186.

ESRI (2001). Building a Geodatabase. ESRI (www.esri.com).

Filion, P., Bunting, T. (1991). Introduction: perspectives on the city. Canadian cities in transition, 2-22. T. Bunting and P. Filion, editors. Toronto: Oxford University Press.

Fowler, M., Scott K. (1997). UML Distilled: A Brief Guide to the Standard Object Modeling Language. Addison-Wesley Longman Inc.

Fransson, U. (1994). Interrelationship between household and housing market: A microsimulation model of household formation among the young. CYBERGEO, No. 135. Institute for Housing Research, Uppsala University, Sweden. (http://193.55.107.3/durham/fransson/fransson.htm).

Fransson, U., Makila, K. (1994). Residential Choice in a Time-Space Perspective: A Micro-Simulation Approach. Neth. J. of Housing and the Built Environment. Vol. 9, No 3.

Goldscheider, C. (1971). Population, modernization, and social structure. Boston: Little, Brown & Co.

Golledge, R.G. (1980). A behavioural view of mobility and migration research. Professional Geographer 32, 14-21.

Goodman, J.L. Jr. (1976). Housing consumption disequilibrium and local residential mobility. Environment and Planning A 8, 855-874.

Hanushek, E.A. and J.M. Quigley (1978a). An explicit model of intra-metropolitan mobility. Land Economics 54, 411-429.

Hanushek, E.A. and J.M. Quigley (1978b). Housing market diseqilibrium and residential mobility. Population mobility and residential change, 51-98. W.A.V. Clark and E.G. Moore, editors. Evanston: Northwestern University.

Harris, R. (1991). Housing. Canadian cities in transition. 350-378. T. Bunting and P. Filion, editors. Toronto: Oxford University Press.

Hong, S., Van den Goor, G.P.M., Brinkkemper, S. (1993). A Formal Approach to the Comparison of Object-Oriented Analysis and Design Methodologies. Proceedings of the 26th Hawaii International Conference on System Sciences, Volume IV, pp. 689-698.

Hoyt, H. (1939). The structure and growth of residential neighbourhoods in American cities. Washington: Federal Housing Administration.

Huff, J.O. and W.A.V. Clark (1978). Cumulative stress and cumulative inertia: a behavioural model of the decision to move. Environment and Planning A 10, 1101-1119.

Jacobson, I., Christerson, M., Jonsson, P., Övergaard, G. (1992). Object-Oriented Software Engineering. Addison-Wesley.

Jackson, R.W. (1994). Object-Oriented Modeling in Regional Science: An Advocacy View. Papers in Regional Science 73, 347-367.

Jensen, C.S., Cliffrod, J., Elmasri, R., Gadia, S.K., Hayes, P., Jacobia, S., (eds), Dyreson, C.E., Grandi, F., Kafer, W., Kline, N., Lorentzos, N., Mitsopoulos, Y., Montnari, A., Nonen, D., Peressi, E., Percini, B., Roddick, J.F., Sadra, N.L., Scalas, M.R., Segev, A., Snodgrass, R.T., Soo, M.D., Tansel, A., Tiberio, R., Wiederhold, G. (1994). A Consensus Glossary of Temporal Database Concepts. SIGMOD Record, Vol. 23, No 1, pp. 52-64.

Jensen, C.S., Snodgrass, R.T. (1997). Temporal Data Management. Technical Report TR-17, TIMECENTER. (Available from http://www.cs.auc.dk/general/DBS/tdb/TimeCenter/publications2.html).

Kain, J.F., Apgar Jr., W.C. (1985). Housing and neighborhood dynamics: a simulation study. Cambridge: Harvard University Press.

Kakoudakis, I. (1996). The TAU Temporal Object Model. MPhil Thesis, UMIST, Manchester, UK.

Kendig. H.L. (1984). Housing careers, life cycle and residential mobility: implications for the housing market. Urban Studies 21, 271-283.

Knox, P.L. (1987). Urban social geography: an introduction. 2nd edition. Essex: Longman Group.

McGinnis, R. (1968). A stochastic model of social mobility. American Sociological Review 33, 712-722.

Michelson, W. (1977). Environmental choice, human behaviour, and residential satisfaction. New York: Oxford University Press.

Miller, E.J., Noehammer, P.J., Ross, D.R. (1987). A Micro-Simulation Model of Residential Mobility. Proceedings of the International Symposium on Transport, Communication and Urban Form: 2, Analytical Techniques and Case Studies, W. Young, ed., Clayton, Victoria: Monash University, pp. 217-234.

Miller, E.J. (1996). Microsimulation and Activity – Based Forecasting. Activity – Based Travel Forecasting Conference Proceedings, June 2-5.

Miller, E.J., Salvini P.A. (2002). Activity-Based Travel Behavior Modeling in a Microsimulation Framework. Invited resource paper, Chapter 26, in H.S. Mahmassani (ed.) In Perpetual Motion, Travel Behavior Research

Opportunities and Application Challenges, Amsterdam: Pergamon, pp. 533-558.

Miller, E.J., Hunt, J.D., Abraham, J.E., Salvini, P.A. (2004). Microsimulating urban systems. Computers, Environment and Urban Systems 28. 9-44.

Moeckel, R., Schurmann, C., Wegener, M. (2002). Microsimulation of Urban Land Use. 42nd European Congress on the Regional Science Association, Dortmund, August 27-31.

Monarchi, D.E., Puhr, G. I. (1992). A Research Typology for Object-Oriented Analysis and Design. Communications of the ACM, Vol. 35, No. 9, September, pp. 35-47.

Moore, E.G. (1972). Residential mobility in the city. Association of American Geographers Resource Paper No. 13. Washington, D.C.: Commission on College Geography.

Newell, C. (1988). *Methods and Models in Demography.* Guilford Pubn Publisher. ISBN-10: 0898627834.

The Object Agency, Inc. (1995). A Comparison of Object-Oriented Development Methodologies. (Available from http://monika.univ.gda.pl/~mattb/objects/TOA/OOMethod/mcr.html).

Orcutt, G.H. (1957). A new type of socio-economic system. Review of Economics and Statistics 39, 116-123.

Orcutt, G.H., H.W. Watts and J.B. Edwards (1968). Data aggregation and information loss. American Economic Review 58, 773-787.

Oskamp, A. (1995). A Microsimulation Approach to Household and Housing Market Modeling. A paper presented to the 1995 Annual Meeting of the American Association of Geographers, Chicago, March 15-18, PDOD Paper No. 29.

Oskamp, A. (1997). Local Housing Market Simulation: A Micro Approach. NETHURD Publications, Amsterdam.

Pfoser, D., Tryfona, N. (1998). Requirements, Definitions and Notations for Spatiotemporal Application Environment. 6th ACM International Workshop on Advances in Geographic Information Systems, ACM Press Baltimore, MD.

Pickvance, C.G. (1974). Life cycle, housing tenure and residential mobility: a path analytic approach. Urban Studies 11, 171-188.

Price, R., Tryfona, N., Jensen, C. S. (1999a). A Conceptual Modeling Language for Spatiotemporal Applications, Chorochronos Technical Report.

Price, R., Srinivasan, B., Kotagiri, R. (1999b). Extending the Unified Modeling Language to Support Spatiotemporal Applications. Asia TOOLS'99, 163-174.

Price, R., Kotagiri, R., Srinivasan, B. (1999c). Spatiotemporal Extensions to Unified Modeling Language. Workshop on Spatio-Temporal Data Models and Languages. Florence, Italy, August 30-31.

Price, R., Tryfona, N., Jensen, C. S. (2000). Extended Spatiotemporal UML: motivations, requirements, and constructs. Journal of Database Management, to be published in Oct/Dec 2000 issue.

Rossi, P.H. (1955). Why families move: a study in the social psycholody of urban residential mobility. Glencoe: The Free Press.

Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F., Lorenzen, W. (1991). Object-Oriented Modeling and Design. Prentice Hall.

Ryan, J., Maoh, H., Kanaroglou, P. (2006). Population Synthesis: Comparing the Major Techniques Using a Small, Complete Population of Firms. Report for the Center for Spatial Analysis (CSpA), School of Geography and Earth Sciences, McMaster University, Hamilton, Ontario, Canada.

Short, J.R. (1978). Residential mobility. Progress in Human Geography 2, 419-447.

Simmons, J.W. (1968). Changing residence in the city: a review of intra-urban mobility. Geographical Review 58, 622-651.

Snodgrass, R., (ed) (1995). The TSQL2 Temporal Query Language. Kluwer Academic Publishers.

Speare, A. Jr., Goldscheider, F.K. (1987). Effects of marital status change on residential mobility. Journal of Marriage and the Family 49, 455-464.

Speare, A. Jr. (1974). Residential satisfaction as an intervening variable in residential mobility. Demography 11, 173-188.

St. John, C., M. Edwards and D. Wenk (1995). Racial differences in intraurban residential mobility. Urban Affairs Review 30, 709-729.

Svinterikou, M. (1998). TUML: A Method for Modelling Temporal Information Systems. MPhil Thesis, UMIST, Manchester, UK.

Theodoulidis, C. I., Loucopoulos, P. (1991). The Time Dimension in Conceptual Modelling. Information Systems, vol. 16, no. 3, pp. 273-300.

Timmermans, H. (2003). The Saga of Integrated Land Use-Transport Modeling: How Many More Dreams Before We Wake Up? 10th International Conference on Travel Behavior Research, Lucerne, 10-15 August.

Tiglao, N. C.C., Tsutsumi, M. (2005). Modeling Households and Location Choices in Metro Manila Using Spatial Microsimulation Approach. Proceedings of the Eastersn Asia Society for Transportation Studies. Vol. 5, pp. 2179-2194.

Tryfona, N., Pfoser, D., Hadzilacos, T. (1997). Modeling Behavior of Geographic Objects: An Experience with the Object Modeling Technique. Proceedings of the 9th International Conference on Advanced Information Systems Engineering. Pages: 347 - 359 , ISBN:3-540-63107-0.

Tryfona, N., Jensen, C. S. (1999). Conceptual Data Modeling for Spatiotemporal Applications, Geoinformatica.

Varady, D.P. (1983). Determinants of residential mobility decisions: the role of government services in relation to other factors. Journal of the American Planning Association 49, 184-199.

Vencatasawmy, C.P., *et al.* (1999). Building a spatial microsimulation model. Paper presented at the 11th European Colloquium on Quantitative and Theoretical Geography in Durham, England, September 3-7.

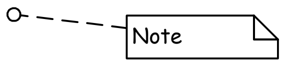
Voas, D., Williamson, P. (2001). Evaluating Goodness-of-Fit Measures for Synthetic Microdata. Geographical & Environmental Modeling, Vol. 5, No. 2, 2001, 177-200.

Waddell, P. (1998). An Urban Simulation model for Integrated Policy Analysis and Planning: Residential Location and Housing Market Components of UrbanSim. 8th World Conference on Transport Research, Antwerp, Belgium, July 12-17.

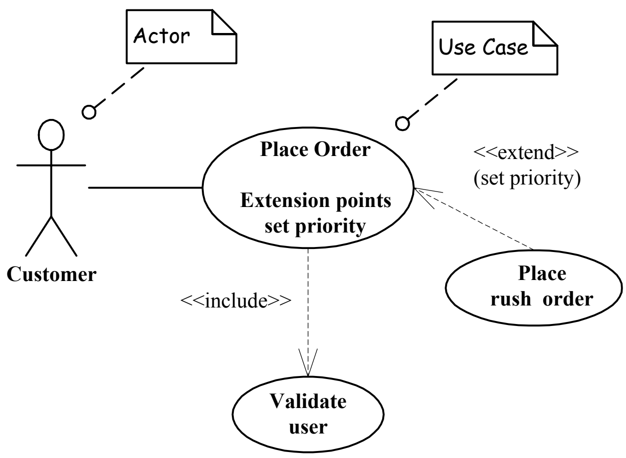
Weeks, J.R. (2001). Population: An Introduction to Concepts and Issues. Wadsworth Publishing. ISBN: 0534529763.

Appendix A
The UML notation

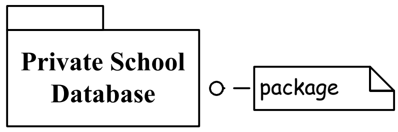
This appendix presents the graphical notation of the elements used in UML diagrams through several examples. Some elements may be used in more than one diagram. The note UML element is used to illuminate element notation.

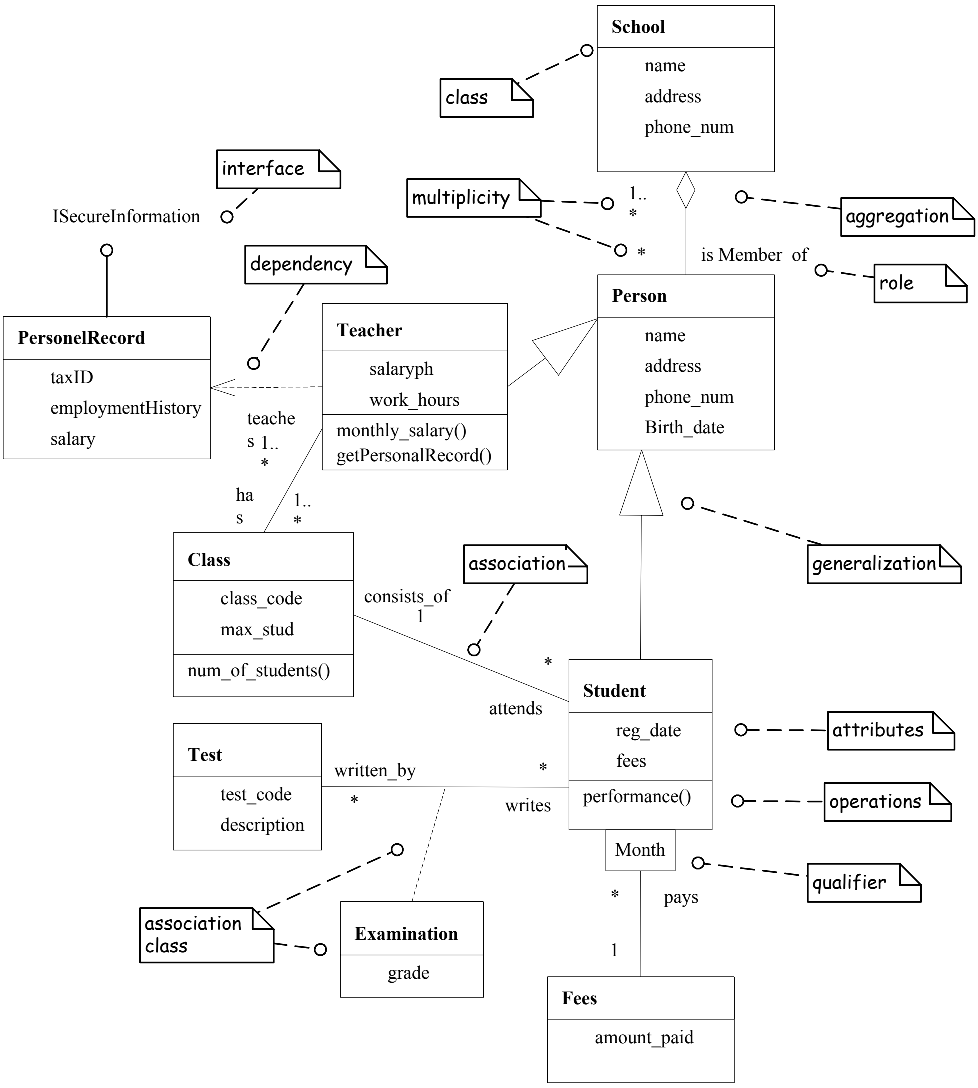


A.1 Use Case Diagram Notation

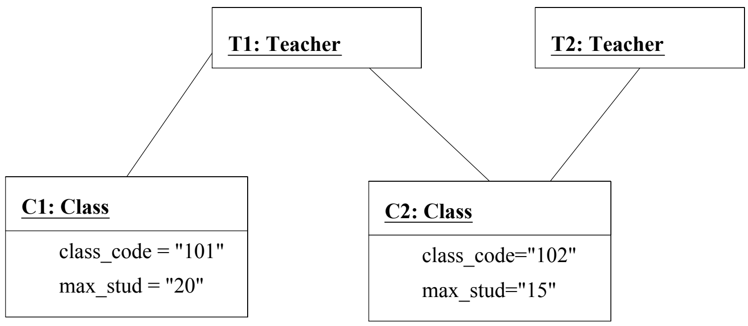


A.2 Class Diagram Notation

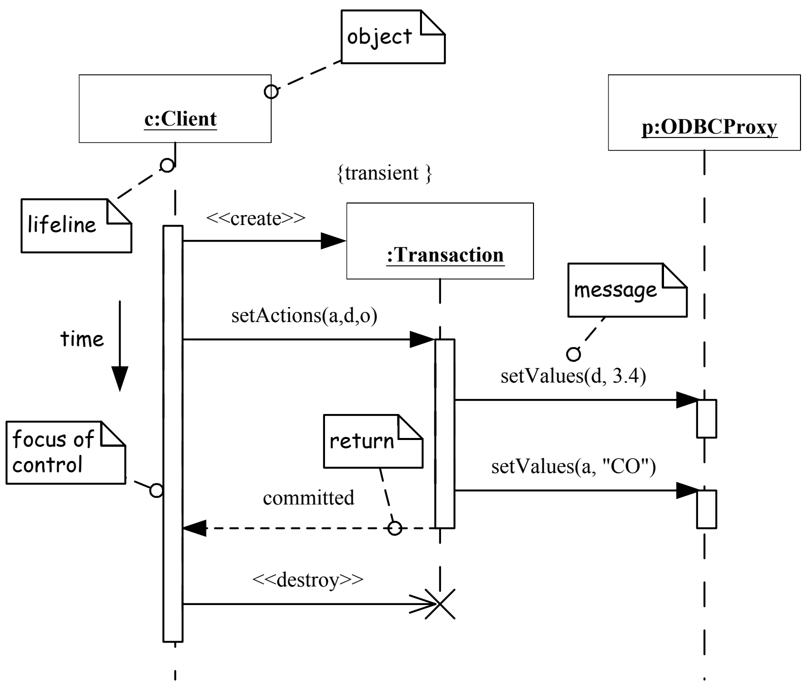




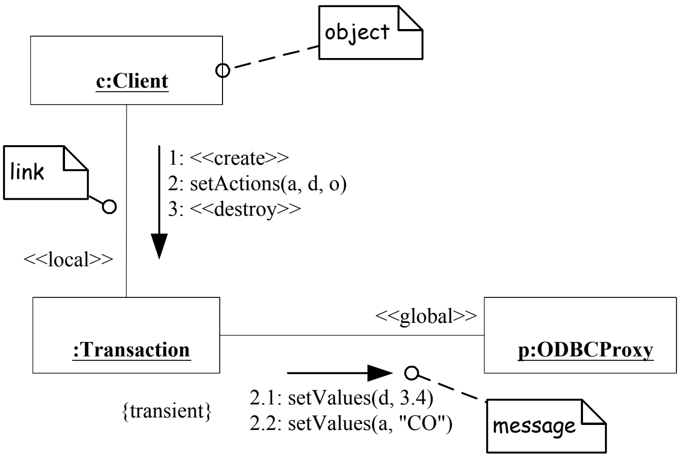
A.3 Object Diagram Notation



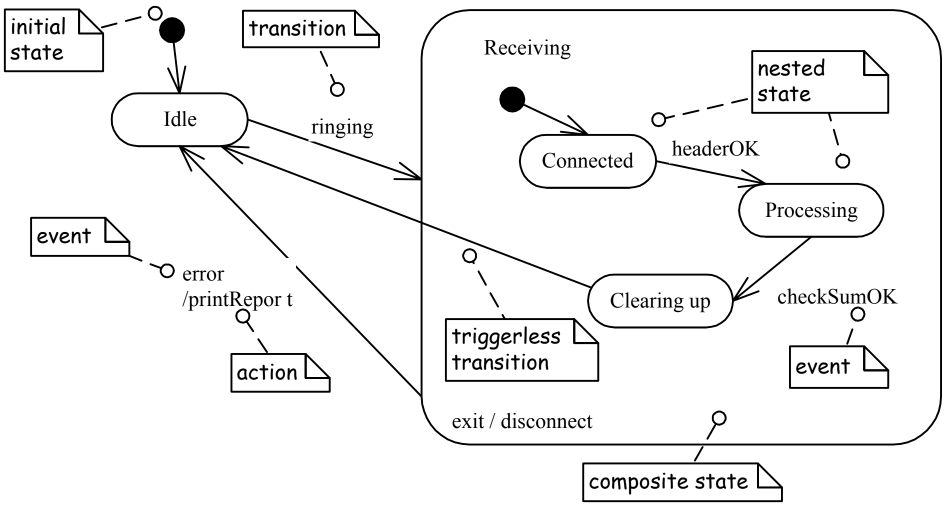
A.4 Sequence Diagram Notation



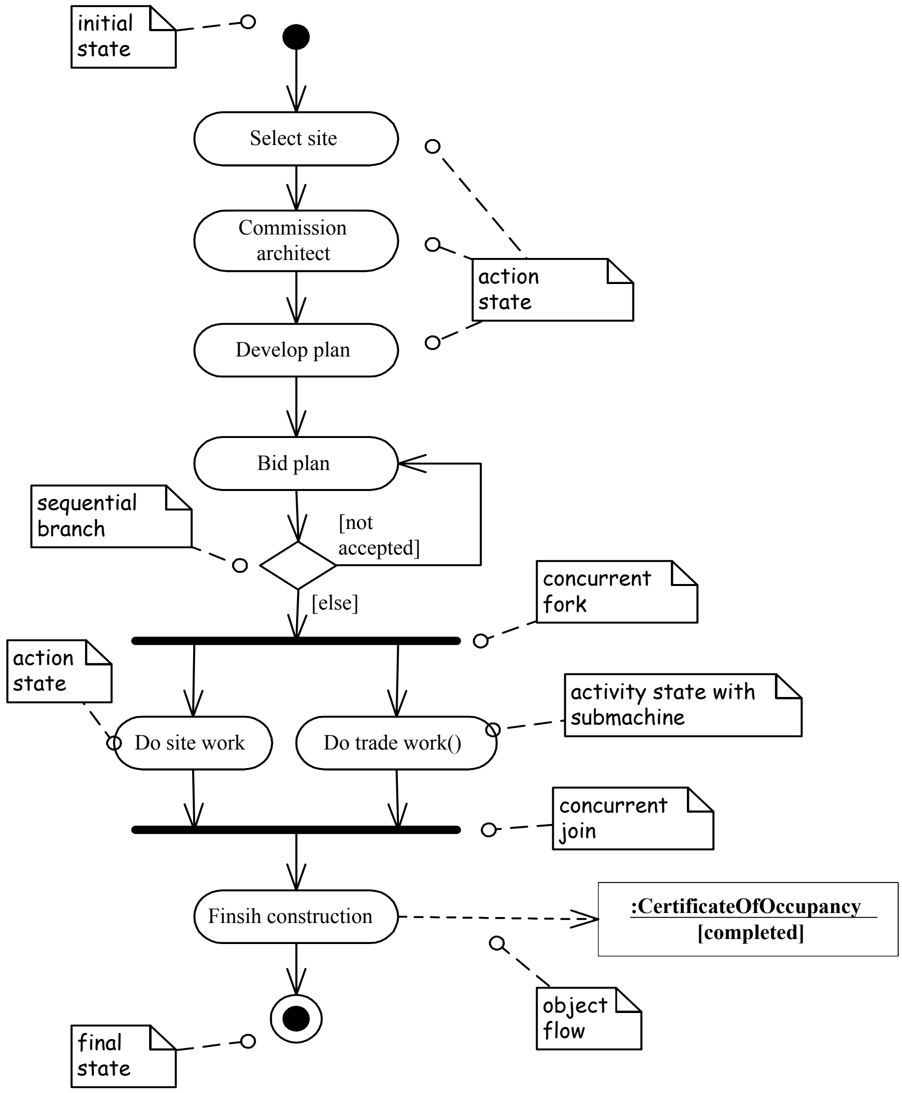
A.5 Collaboration Diagram Notation



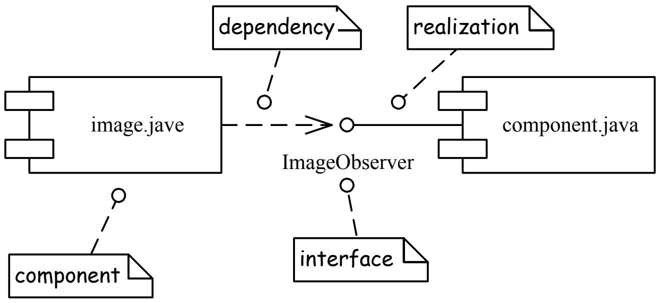
A.6 Statechart Diagram Notation



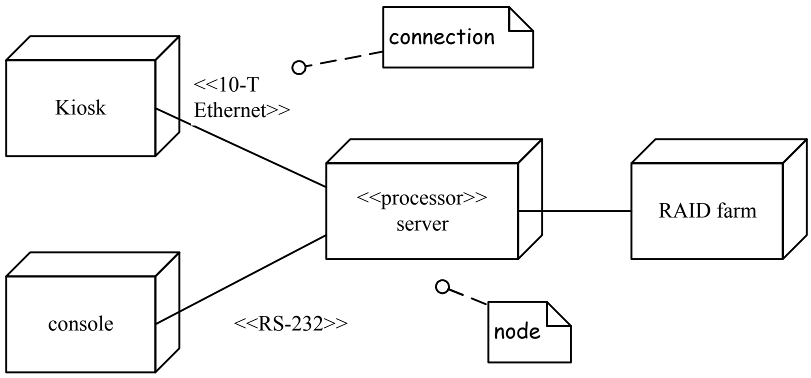
A.7 Activity Diagram Notation



A.8 Component Diagram Notation



A.9 Deployment Diagrams Notation



Appendix B
The STUML process: A User's Guide

This appendix presents the steps of the STUML process in detail. The STUML process is use-case driven, architecture-centric, iterative and incremental. It consists of the macro and micro development processes.

B.1 The Macro development process

The macro process is the framework for controlling the micro process. During the macro process a set of activities takes place that leads to a number of products and allows the development team to identify risks and make early corrections to the micro process. It represents the activities of the development team during the entire development life cycle. The high-level activities that comprise the macro process are:

1. *Conceptualisation*: establish the core requirements for the software.
2. *Analysis*: develop a model of the system's desired behaviour.
3. *Design*: create an architecture for the implementation.
4. *Evolution*: evolve the implementation though successive refinement.
5. *Maintenance*: manage post-delivery evolution.

The macro process repeats itself after major product releases. This particularly applies to the development of families of products. Next, a description of the macro process activities is given.

B.1.1 Conceptualisation

The purpose of conceptualisation is to identify the core requirements for the system to be developed, to define the scope of the system and to prove that the system may actually be developed. The primary products of this activity are prototypes. Each prototype implements a particular system development path with specific assumptions concerning functionality, performance, size, complexity, etc. The prototypes serve as the proof of concept and their evaluation reveals whether the potential risks permit the system development. The prototypes developed during this activity are by their nature incomplete and should be put aside once the activity of conceptualisation is completed (Booch, 1994).

B.1.2 Analysis

The purpose of analysis is to define the function of the system, that is, to determine what the system does. This is achieved through *domain analysis* and *use case planning*.

Domain analysis (Booch, 1994) is the examination of existing related systems in order to identify the classes and objects that are common to all applications within a given domain such as patient record tracking, compilers, etc. This activity helps the development team to identify the behaviour of the system under development, as it indicates the classes and objects that have proven useful in other related systems. Moreover, the development team benefits from the experience of related system development and there are cases where domain analysis leads to the conclusion that existing software may be reused or adapted and therefore, there is no need for the development of new software.

Use case planning is the central activity of analysis. The order of events that takes place during use case planning is:

- Identify all the primary use cases of the system, namely all the primary activities. If possible, group them into sets of use cases with related functionality and into use case hierarchies where high level use cases build upon more primitive ones.
- Elaborate on each interesting use case and determine the associated primary scenarios. A use case scenario refers to a single path through a use case (Fowler and Scott, 1997) and serves as the starting point for defining the semantics of the use case.
- Define the semantics of each use case. In particular, use interaction diagrams (collaboration and/or sequence diagrams) to show how objects collaborate to carry out the activities of the use case. Additionally, use class diagrams to show the semantics of the objects that take part to the use case. Finally, if the use case contains significant workflow elements, use activity diagrams to capture them. Document any assumptions, constraints, or performance issues for each use case.
- If necessary, elaborate on use cases to show behaviour under exceptional conditions.
- If the life cycle of certain objects is complex or significant, use a statechart diagram for the class of objects to describe it.
- Search for patterns among use cases and express them in terms of class diagrams to show the similarities among different use cases.
- Update the data dictionary with the classes and objects identified in the previous events.

The activity of analysis is completed when all the primary use cases have been developed and validated as well as the set of secondary use cases that are considered to be important to the system's scope. The evaluation of the diagrams produced during analysis reveals the potential risks that may impact the design activity.

B.1.3 Design

The purpose of design is to define the architecture of the system under development and to determine the common policies that must be used by different parts of the system. This is achieved through the activities of *architectural planning*, *tactical design* and *release planning*.

According to (Booch, 1994) the architecture of an OO software system encompasses its class and object structure, organised in terms of distinct layers and partitions. Therefore, *architectural planning* involves the creation of the layers and partitions that comprise the architecture of the system. During this activity, the packages and deployment diagrams are particularly useful. Packages show the logical pieces of the system under development and their dependencies. Deployment diagrams show the physical relationships among software and hardware components of the delivered system. The order of events that takes place during this activity is:

- Consider grouping the use cases identified during analysis, and allocate these to layers and partitions of the architecture. Use cases that build upon one another should fall into different layers; use cases that collaborate at a similar level of abstraction should fall into partitions of the same layer.
- Develop an executable release that implements parts of some use cases in order to validate the architecture.
- Evaluate the architecture and estimate its weaknesses and strengths. Identify the potential risks of the architecture so that resources can be accordingly allocated at the beginning of evolution.

Architectural planning aims to create a domain-specific application framework that may be successively refined, early in the development process.

Tactical design involves the design of mechanisms that appear throughout the system. Examples of such mechanisms are the policies for error detection and handling, memory management etc. Some policies address domain-independent issues such as memory management and error handling while others address domain-specific issues such as transaction and database management in information systems. The order of events that take place is:

- Determine the common policies that must be used by different parts of the system.
- For each policy develop a use case to demonstrate the behaviour of that policy. In addition, develop a prototype that implements it.
- Document each policy.

An issue specific to temporal information systems is the *longevity of data*, that is, the time period that information referring to past states of the application objects is kept in the temporal database. The physical deletion of past information is a desirable feature of a temporal application (Snodgrass, R., (ed), 1995). That is why, during tactical design the development team should design the policy concerning the longevity of data. Initially, they should decide when data should be removed from the database: after a fixed period of storage in the database, the decision is left to the database administrator etc. At this stage, the team should design the mechanism for the physical deletion of the data.

During *release planning* the plan for the development of a series of software releases is defined. Each release will incorporate more functionality than the previous one aiming to ultimately encompass the requirements of the entire system. The plan is based on the use cases and the risk assessment generated during analysis. The order of events that takes place during this activity is:

- Define the level of priority for each use case identified during analysis.

- Allocate use cases to a series of software releases whose final delivery will be the production system.
- Set the goals and schedules of the releases and ensure that delivery dates allow adequate development time and that releases are synchronised with other activities such as documentation.
- Begin task planning in order to identify the work and the resources that are necessary for each release.

The activity of design is completed when a description of an architecture, a description of the common policies and a plan for successive releases has been produced.

B.1.4 Evolution

The purpose of evolution is to construct the system in a series of releases, wherein each release is a refinement of the previous one. These releases are driven by the plan set in release planning. At this phase constraints such as functionality, time and space should be taken into consideration in order to produce a system that will actually serve the purposes of its development. The work that is carried out between two releases is a compressed development process. The development team performs analysis, design and implementation for the use cases assigned to each release. The order of events that takes place is:

- Identify the use cases that should be implemented by this release and the areas of highest risk, based on the evaluation of the previous release and the release plan.
- Assign tasks to the team to carry out this release. Based on the products of the analysis phase and the products of the previous release, the team should perform the activities of micro process in order to identify the classes and objects necessary to implement the use cases. During this

task, changes may occur in existing classes and these changes should be controlled in order to prevent the destruction of the system architecture. In (Booch, 1994) this is referred to as the activity of change management. During the evolution of a system the kinds of changes to be expected are (Booch, 1994):
  - Adding a new class or a new group of classes;
  - Changing the implementation of a class;
  - Changing the representation of a class;
  - Reorganising the class structure;
  - Changing the interface of a class.

  When a change in a class occurs consideration must be given to the way this change affects other classes.
- Assign developers to produce behavioural prototypes. A behavioural prototype serves to explore design alternatives (such as a new algorithm) so that areas of risk may be resolved early.
- Produce the product release.
- Reconsider the original release plan and make the required adjustments for subsequent releases.

The phase of evolution is completed when the functionality and quality of the release meet the user requirements.

B.1.5 Maintenance

Maintenance is in fact a continuation of the evolution phase, during which more localised changes are made to the system (through successive releases) as new requirements are added and bugs are being detected. The activities in this phase are the same as the ones that take place during the previous phase. In addition, during maintenance the *planning* activity takes place, which prioritises the tasks to be completed.

B.2 The Micro development process

The micro process is driven by the set of use cases and product releases that result from, and are successively refined by, the macro process. It mainly represents the daily activities of the individual developer or a small team of developers. In the micro process the traditional phases of analysis and design are intentionally blurred. The activities that comprise the micro process are:

1. *Identify the classes and objects at a given level of abstraction.*
2. *Identify the semantics of these classes and objects.*
3. *Identify the relationships among these classes and objects.*
4. *Identify the spatial semantics of these classes, objects and relationships.*
5. *Identify the temporal semantics of these classes, objects and relationships.*
6. *Specify the interface and then the implementation of these classes and objects.*

B.2.1 Identify classes and objects

The purpose of this activity is to establish the boundaries of the system under development. The starting point for identifying classes and objects is the use cases and the use case scenarios identified during the macro process. The development team walks through each use case and each use case scenario in order to identify the objects that participate in each one, and to determine their responsibilities and how these objects collaborate to carry out the activities of the use case.

During analysis, this step reveals the classes and objects that form the vocabulary of the system domain. During design, this step reveals new classes and objects that should be part of the system under development.

During implementation, this step discovers lower-level classes and objects that may be used to construct higher-level ones.

As the development team learns more about the system under development, some of the classes identified early in the lifecycle are discarded, while others carry through all the way to implementation. Moreover, the boundaries of certain classes change by reallocating responsibilities, splitting larger classes to groups of collaborating ones, etc.

The product of this step is a data dictionary of all significant classes that is updated as development proceeds. The activity of identifying classes and objects is completed when a stable data dictionary is produced that contains an adequate set of classes, consistently named and with a reasonable separation of responsibilities.

B.2.2 Identify the semantics of classes and objects

The purpose of this activity is to identify the responsibilities, that is, the behaviour and attributes of each class identified in the previous phase. During analysis, this step identifies the responsibilities of the classes derived from the system domain. During design, this step distributes the responsibilities among the parts of the system under development. During implementation, this step specifies the interface of each class and defines the concrete signature of each operation. The activities that take place during this phase are the *responsibility definition,* the *isolated class design* and the *pattern search*.

The starting point of the *responsibility definition* is the use cases and the use case scenarios generated by the macro process. The order of events that takes place during this activity is:

- For each use case and use case scenario reconsider the classes and objects identified in the previous phase.
- Walk through each use case and each use case scenario, allocating responsibilities to each object sufficient to accomplish the desired behaviour. Use interaction diagrams, namely collaboration and sequence diagrams, to show the distribution of responsibilities among the objects of each use case. Moreover, use statechart diagrams to show the responsibilities of certain classes whose behaviour depends on their current state.
- Reallocate responsibilities so that there is a balanced distribution of behaviour. Consider reusing responsibilities, splitting large responsibilities into smaller ones and combining trivial responsibilities to larger ones.

Early in the development process the responsibilities of each class and each object may be defined in natural language. As the development process proceeds, the names of the operations that satisfy each responsibility are defined, and finally the full signature of each operation is specified.

During *isolated class design*, the development team focuses on a single class and defines the set of operations that satisfy the responsibilities of the class. The order of events that takes place during this activity is:

- Select one class and identify its responsibilities.
- Specify a set of operations that satisfy these responsibilities. Consider reusing operations for similar responsibilities.
- Consider each operation and check whether it is primitive. If not, split it into more primitive ones. Composite operations may be kept in the class itself or defined in a utility class.

- Consider the policy for construction, copying, and destruction. Follow a common policy for these behaviours and do not allow the class to have its own idiom unless that is absolutely necessary. This event takes place later in the development process.
- Consider the need for completeness; add other primitive operations that complete the semantics of the class although they are not required by the present clients.

If the class is part of an inheritance lattice, then consider placing the operations of the class higher in the hierarchy and consider introducing new classes if necessary. For example, operations that may be used by two or more peer classes should be placed in a common superclass.

During pattern search, the development team searches for common behaviours in order to exploit opportunities for reuse. The order of events that takes place during this activity is:

- Search the set of use cases and use case scenarios for patterns of collaborations among classes/objects and explicitly document the non-trivial ones so that they can be reused rather than reinvented.
- Search the set of responsibilities for patterns of behaviour and consider unifying common responsibilities, for example, in common abstract classes.
- Search for patterns within operation signatures. This event takes place later in the development process when operation signatures are specified. Remove any unnecessary differences and introduce, for example, utility classes that contain operations that are found in two or more classes.

The products of this activity are a refined data dictionary that includes the responsibilities assigned to each class and each object as well as a set of collaboration diagrams and sequence diagrams that capture the semantics of the use cases. The activity of identifying the semantics of classes and objects is completed when an adequate, primitive and complete set of responsibilities and operations is produced for each class and each object.

B.2.3 Identify the relationships among classes and objects.

The purpose of this activity is to clearly define the boundaries of each class and to identify the classes that collaborate with each class. The activities that take place during this phase are the *specification of associations*, the *identification of various collaborations* and the *refinements of associations*.

The *specification of associations* is mainly an analysis and early design activity during which the development team captures the details of the relationships between classes without having to make premature design decisions. The order of events that takes place during this activity is:

- Select the classes of one or more use cases, and populate the diagram with the operations and attributes of each class as needed to show the significant properties of the system part being modelled.
- Consider the presence of a semantic dependency between any two classes such as the need for navigation from one object to the other and the need for triggering some behaviour from an object, and define a binary association if such a dependency exists. Next, consider the presence of a semantic dependency between three or more classes and define an n-ary association if such a dependency exists. This kind of association is not often encountered and it should be used only when it cannot be replaced by a semantically equivalent set of binary associations. Some associations may immediately be identified as

generalisation relationships, aggregation associations or composition associations.

- For each association specify the role and the multiplicity of each participating class as well as any other constraint that applies to the association ends or to the association itself.
- Validate these decisions by walking through use cases and use case scenarios, and verifying that associations provide the navigation and behaviour required by each use case.

The *identification of collaborations* is mainly a design decision. The development team should consider several kinds of collaborations, depending on the phase of the macro process in which the team is:

- Specify the mechanisms identified in the previous step through object diagrams, each illustrating the dynamic semantics of a mechanism. Validate each mechanism by walking through use case scenarios and consider introducing concurrency. During this activity, new paths may be introduced among objects or existing ones may be removed.
- Consider placing the classes into a generalisation/specialisation hierarchy. Where necessary introduce new classes as part of the hierarchy. Review the resulting class diagram in order to ensure that the hierarchy lattice is well balanced, that is, it is not too tall or too short and neither too wide nor too shinny. If patterns of structure or behaviour appear among classes, reorganise the hierarchy only if there is no loss of simplicity.
- Consider grouping classes into packages and components into nodes. This is done by taking a global view of the system and splitting it into parts. Each part should denote a service offered by the system, which is logically cohesive and/or likely to change independently from the others.

- Consider allocating classes and objects to components. This activity is part of mapping the system model to actual code and enables the development team to make decisions on the implementation of the classes identified so far.

The *refinement of associations* is both an analysis and design activity. During analysis some of the associations are semantically refined as a result of the developments team's increasing understanding of the system under development. During design, the associations are refined and even new ones are added in order to be closer to implementation. The order of events that takes place during this activity is:

- Select a set of classes that are grouped by some set of associations and search for patterns of behaviour that indicate the opportunity for introducing generalisation relationships. Place the classes in an existing inheritance lattice, or create a new one if an appropriate lattice does not already exist.
- Search for patterns of structure and capture them by introducing new classes through inheritance or aggregation.
- Search for classes with similar behaviour that are either disjoint peers or not part of an inheritance lattice and consider introducing parameterised classes.
- Consider the navigability of existing associations.
- Consider defining the characteristics of the existing associations such as keys, multiplicity etc.; include only the association characteristics that are important for the analysis or design phase, or that are necessary for implementation.

The products of this step are a refined data dictionary that includes the relationships defined among classes and objects, as well as a set of class, collaboration and deployment diagrams. The activity of identifying the relationships among classes and objects is completed when the relationships between the important classes have been specified.

B.2.4 Identify the spatial semantics of classes, objects and relationships

The purpose of this phase is to define the spatial semantics of the classes and objects that were identified as part of the system domain earlier in the micro process. When considering the spatial semantics of the classes in a hierarchy lattice, the development team should start by considering the root of the hierarchy lattice, and then continue by examining the classes all the way down to the leaves. For each class the order of events that take place is:

- *Identify the spatial semantics of the attributes.* For each attribute of the class consider whether:
  - The attribute value should be a spatial extent (spatial attribute).
  - The attribute should have a set of thematic values, each associated with a spatial extent representing the location where that attribute value is valid, i.e. the attribute values may change over space and their changed values should be retained (spatially dependent thematic attribute).
  - For each spatial and spatially dependent thematic attribute, consider its spatial dimension(s) (0-D, 1-D, 2D).
  - For each attribute that the class inherits from its superclasses, consider changing its spatial dimensions (e.g. 0-D to 2-D) or changing the number of spatial dimensions defined (e.g. 1-D to 1-D and 2-D).
- *Define the spatial semantics of the class.*
  - Consider whether the objects' positions in space are of interest (spatial objects/class).

  - A class with one or more spatial or spatially dependent thematic attributes should be defined as a spatial one, i.e. its objects should be associated with one or more spatial extents.
- *Identify the spatial semantics of the associations.* If the class has spatial semantics, then for each association identified earlier in the micro process in which the class participates, consider whether:
  - Each association instance (link) should be associated with a spatial extent representing the location where the link is valid, i.e. the link may change over space and its changes should be retained (spatially dependent association).
  - For each spatially dependent association consider its spatial dimension (0-D, 1-D, 2D) based on the spatial dimensions of the participating classes.
- *Define the spatial semantics of the classes that belong to the same hierarchy as the class under consideration.* If the class has one or more superclasses with spatial semantics, then the class as well as its subclasses should also have spatial semantics.

Early in the development process, spatial and spatially dependent thematic attributes, spatial classes and spatially dependent associations are defined. As development proceeds, the spatial semantics of classes and their properties are re-examined and are refined to include the spatial dimensions and interpolation of information.

The products of this activity are a refined data dictionary as well as a set of class diagrams that capture the spatial semantics of classes and relationships. The activity of identifying the spatial semantics of classes, objects and relationships is completed when the spatial semantics of all the significant classes and relationships have been specified.

B.2.5 Identify the temporal semantics of classes, objects and relationships

The purpose of this phase is to define the temporal semantics of the classes and objects that were identified as part of the system domain earlier in the micro process. When considering the temporal semantics of the classes in a hierarchy lattice, the development team should start by considering the root of the hierarchy lattice, and then continue by examining the classes all the way down to the leaves. For each class the order of events that take place is:

- *Identify the temporal semantics of the attributes.* For each attribute of the class, its temporal semantics are considered. In particular:
  - For each thematic attribute consider whether the attribute should have a set of thematic values, each associated with one or more timestamps (the value's valid and/or transaction time), i.e. the attribute values may change over time and their changed values should be retained (temporally dependent thematic attribute).
  - For each spatial attribute consider whether its value should be associated with one or more timestamps representing the spatial extent's valid and/or transaction time (temporally dependent spatial attribute).
  - For each spatially dependent thematic attribute consider whether the attribute should have a set of thematic values, each associated with a spatial extent and one or more timestamps (spatio-temporally dependent thematic attribute).
  - For each attribute that the class inherits from its superclasses, consider specialising its temporal semantics. In particular, an attribute without any temporal semantics may be refined to a transaction, valid or bitemporal time attribute, while a transaction or valid time attribute may be refined to a bitemporal time attribute.

- *Identify the temporal semantics of the associations*. For each association identified earlier in the micro process in which the class participates, consider its temporal semantics. In particular:
  - For each association without spatial semantics consider whether the association should have a set of links, each associated with one or more timestamps (the value's valid and/or transaction time), i.e. the links may change over time and their changes should be retained (temporally dependent association).
  - For each spatially dependent association consider whether the association should have a set of links, each associated with a spatial extent and one or more timestamps (spatio-temporally dependent association).
  - For each association that the class inherits from its superclasses, consider specialising its temporal semantics. In particular, an association without temporal semantics may be refined to a transaction, valid or bitemporal time association, while a transaction or valid time association may be refined to a bitemporal time association.
- *Define the temporal semantics of the class.*
  - If the class is not a spatial one, consider whether each object of the class should be associated with one or more timestamps, representing the object's valid and/or transaction time (temporal class).
  - If the class is a spatial one, consider whether:
    - Each object of the class should be associated with a spatial extent and one or more timestamps representing the object's valid and/or transaction time (spatiotemporal class).
    - Each object of the class should have its spatial extent associated with one or more timestamps, representing the spatial extent's valid and/or transaction time (temporally dependent spatial class).

- Each object of the class should be associated with one or more timestamps, representing the object's valid and/or transaction time and should have its spatial extent associated with one or more timestamps, representing the spatial extent's valid and/or transaction time (temporally dependent spatiotemporal class).

The temporal semantics of the class must by default include all the temporal dimensions that its properties have. Moreover, the class may have more temporal dimensions. In particular, a class without temporal semantics may only have non-temporal properties, a valid time class may only have valid time and/or non-temporal properties and a bitemporal time class may have bitemporal time, valid time, transaction time and/or non-temporal properties. For example, a class that has non-temporal properties, one or more valid time properties and one or more transaction time properties, should be defined as a bitemporal time class. On the other hand, a class that has non-temporal properties and one or more valid time properties may be defined as a valid time or a bitemporal time class.

- *Define the temporal semantics of the classes that belong to the same hierarchy as the class under consideration*. If the class has superclasses and/or subclasses, then update the hierarchy lattice so that each subclass either belongs to the same temporal category as its superclass, or has more temporal dimensions. In particular, a non-temporal class should have as a subclass non-temporal, transaction, valid or bitemporal time classes. A transaction time class should have as a subclass transaction or bitemporal time classes. A valid time class should have as a subclass valid or bitemporal time subclasses. Finally, a bitemporal time class may have as a subclass only bitemporal time classes.

Early in the development process, valid, transaction and bitemporal classes, attributes and relationships are defined. As development proceeds, the temporal semantics of classes and their properties are re-examined and are refined to include the granularity and type of the valid time timestamp of each valid and bitemporal time information.

The products of this activity are a refined data dictionary as well as a set of class diagrams that capture the temporal semantics of classes and relationships. The activity of identifying the temporal semantics of classes, objects and relationships is completed when the temporal semantics of all the significant classes and relationships have been specified.

B.2.6 Implement classes and objects

The purpose of this activity is to define the representation of the classes and objects defined in the previous activities, and to define the mapping of these representations to actual code. During analysis, the purpose of this activity is to refine the existing classes and objects in order to reveal new classes and objects to which responsibilities may be delegated. During design the purpose of this activity is to create executable or near-executable representations of the existing classes and objects. The order of events that takes place during this activity is:

- For each class consider its interface. Identify the operations that should be optimised and define the signatures of all the important operations.
- Before choosing a representation from scratch, consider the use of inheritance for implementation, or the use of parameterised classes.
- Consider the objects to which responsibilities may be delegated.
- If the class' semantics can not be provided through the representations mentioned in the two previous events, consider a suitable representation from scratch.

- Select a suitable algorithm for each operation. Divide complex algorithms into a set of auxiliary operations that are less complicated and more likely to be reused.

The products of this activity are a refined data dictionary as well as a set of diagrams (class, interaction, statechart and/or deployment diagrams).

Appendix C
The UML use case scenarios

This appendix presents the *use-cases scenarios* that define in detail the events modeled by *URM-Microsim.*

C.1 The Housing Demand use case scenarios

The Housing Demand sub-model consists of the following use cases:

C.1.1 Demographic change use case scenario

Use Case Name: Demographic change

Summary: The demographic changes and the development of the population are determined.

Actor: Household, Individual

Precondition: -

Description:

- Each individual is subjected to the probability of experiencing one of the *death*, *fertility, union formation, union dissolution, flat-mate formation, flat-mate leaving, nest leaving* events and *out-migration*, based on the individual's current life cycle stage, as well as the previous event experienced by the individual's household.
- The *In-migration* use case is called.

Postcondition: The household denotes whether it intends to search for a new house or not.

C.1.2 Death use case scenario

Use Case Name: Death

Summary: An individual dies.

Actor: Individual

Precondition: An individual is subjected to a probability of dying based on his/her age and sex.

Description:

1. The individual dies.
2. The individual is removed from the population.
3. If the individual was a worker then his/her job is denoted as available.
4. The number of deaths increases by one.
5. The household size is decreased by one.
6. The household expresses the intention to move to another dwelling

Alternatives:

- If the individual belonged to a one-member household, then:
  5. The household is removed from the household set.
  6. The dwelling the household occupied is vacated and added to the housing supply set.
- If the individual was parent in a one-parent household with one or more children, then consider defining a guardian (>=18 years old) from the remaining household members. If no guardian is available then:
  5. The household is removed from the household set.
  6. The children are added to the orphan set.
  7. The dwelling the household occupied is vacated and added to the housing supply set.

Postcondition: The household intends to search for a new house with *low intensity*.

C.1.3 Fertility use case scenario

Use Case Name: Fertility

Summary: An individual gives birth to one or more individuals

Actor: Individual

Precondition: A woman of age >= 10 is subjected to a probability of giving birth based on her age.

Description:

1. The individual gives birth to a child.
2. A new individual is created.
3. The sex of the child is determined.
4. The child is subjected to the infant mortality probability.
5. If the child survives, it becomes a member of the household to which the mother belongs.
6. The household size increases by one.
7. The household expresses the intention to move to another dwelling.

Alternatives:

- If more than one child is born, the scenario is repeated for each child.
- If the infant dies, then the number of deaths increases by one and the use case terminates.

Postcondition: The household intends to search for a new house with *low intensity*.

C.1.4 Union formation use case scenario

Use Case Name: Union formation

Summary: An individual finds a partner and decides to live in with her/him.

Actor: Individual

Precondition: An individual, who is of age >= 10 and not married, is searching for a partner.

Description:

1. The individual searches for a not married individual of the opposite sex and of certain age.
2. If the individual finds a partner, then the couple decides to live together (form a household).
3. If both individuals live with their parents or flat-mate, then a new household is created.
4. The new household expresses the intention to find a dwelling.

Alternatives:

- If the individual does not find a partner, then the use case terminates.
- If one partner already has a dwelling of his/her own, then:
  3. The other partner moves in with him/her.
  4. The existing household is updated and its size is increases by one.
  5. The household, to which the moving partner belonged, decreases in size by one
  6. The household expresses the intention to search for a house with low intensity.
- If both partners already have a dwelling of their own, then:
  3. The couple decides which dwelling to keep.
  4. The corresponding household is updated and its size is increases by one.
  5. The household of the moving partner is removed from the household set.
  6. Its dwelling is vacated and added to the housing supply set.
  7. The household expresses the intention to search for a house with low intensity.

Postcondition: The new household intends to search for a house with *medium intensity*.

C.1.5 Union dissolution use case scenario

Use Case Name: Union dissolution

Summary: A couple that belongs to the same household (married or cohabitates) decides to separate.

Actor: Household

Precondition: A couple is subjected to a probability of splitting based on the female's age and the years of marriage.

Description:

1. The household is simulated to split.
2. It is decided which parent keeps the children (if any).
3. It is decided which partner (male or female) moves out of the current dwelling.
4. A new household is created and consists of the partner that moves out.
5. The original household size is decreased by the number of moving out members.
6. The new household must move to another dwelling.

Alternatives:

- If the new household decides to leave the housing market under examination, then the use case terminates.

Postcondition: The household intends to search for a new house with *high intensity*.

C.1.6 Nest leaving use case scenario

Use Case Name: Nest leaving

Summary: An individual, who lives with his/her parents, decides to leave.

Actor: Individual

Precondition: An individual of age >= 18, who still lives with his/her parents is subjected to a probability of moving out.

Description:

1. The individual is simulated to leave his/her parents.

2. The original household decreases in size.
3. If the individual decides to remain in the housing market under examination (the individual should have a *job/income*), then he/she forms a new single-member household.
4. The new household expresses the intention to move to another dwelling.

Alternatives:

- If the individual decides to leave the housing market under examination (to search for a *job* or study at a *university*), then the use case terminates.

Postcondition: The new household intends to search for a new house with *medium intensity*.

C.1.7 Flat-mate formation use case scenario

Use Case Name: Flat-mate formation

Summary: An individual who lives on his/her own or flat-mates with others decides to find another individual to flat-mate.

Actor: Individual

Precondition: An individual who lives on his/her own or flat-mates with others is subjected to the probability of finding another flat-mate.

Description:

1. The individual searches for an individual of the same sex and age range who lives on his/her own or flat-mates with others.
2. If the individual finds a flat-mate, then they decide to flat-mate (form a household).
3. If both individuals live with their others, then a new household is created.
4. The new household expresses the intention to find a dwelling.

Alternatives:

- If the individual does not find a flat-mate, then the use case terminates.
- If one flat-mate already has a dwelling of his/her own, then:

3. The other flat-mate moves in with him/her.
4. The existing household is updated and its size is increased by one.
5. The household, to which the moving flat-mate belonged, decreases in size by one
6. The household expresses the intention to search for a house with low intensity.

- If both flat-mates already have a dwelling of their own, then:
  3. They decide which dwelling to keep.
  4. The corresponding household is updated and its size is increased by one.
  5. The household of the moving flat-mate is removed from the household set.
  6. Its dwelling is vacated and added to the housing supply set.
  7. The household expresses the intention to search for a house with low intensity.

Postcondition: The new household intends to search for a house with *medium intensity*.

C.1.8 Flat-mate leaving use case scenario

Use Case Name: Flat-mate leaving

Summary: An individual, who flat-mates with others, decides to leave.

Actor: Individual

Precondition: An individual who flat-mates with other individuals is subjected to the probability of leaving the household to live on his/her own.

Description:

1. The individual is simulated to leave the household.
2. The original household decreases in size.
3. If the individual decides to remain in the housing market under examination, then he/she forms a new single-member household.
4. The new household expresses the intention to move to another dwelling.

Alternatives:

- If the individual decides to leave the housing market under examination, then the use case terminates.

Postcondition: The new household intends to search for a new house with *medium intensity*.

C.1.9 Out-migration use case scenario

Use Case Name: Out-migration.

Summary: An individual or a household decides to migrate out of the housing market.

Actor: Individual, Household

Precondition:

1. A child of age 18 and above may decide to study at the University of another city.
2. If a household searches for a dwelling with high intensity and fails to find one, then the household is subjected to the probability of out-migrating based on the years of housing search.
3. If the primary worker of a household is unemployed or cannot find a job that satisfies him/her, then the household is subjected to the probability of out-migrating based on the years of unemployment.

Description:

1. The household is simulated to out-migrate.
2. The household is removed from the household set and its members from the population.
3. If one or more members of the household were workers, then their jobs are denoted as *available*.
4. The household's dwelling is vacated and added to the housing supply set.

Alternatives:

- If only one individual is simulated to out-migrate then the original household decreases in size.

Postcondition: The housing market is updated.

C.1.10 In-migration use case scenario

Use Case Name: In-migration

Summary: Households outside the housing market decide to in-migrate.

Actor: Household

Precondition: The number of households entering the housing market is determined.

Description:

1. A new household is formed for each in-migrating household.
2. The individuals – members of each household – are created.
3. The new households must find a dwelling.

Postcondition: The new households intend to search for a new house with *high intensity*.

C.1.11 Income change use case scenario

Use Case Name: Income change

Summary: An individual has its income changed due to its occupation.

Actor: Individual

Precondition: An individual of age 18 and above.

Description:

- If the education career of the individual is not already set then the *Education* use case is called.
- If the individual is not a worker, then the *Job entry* use case is called.
- If the individual already works, then the *Job change* use case is called.

Alternatives:

- The worker becomes unemployed and his/her salary is set to the unemployment allowance. It is decided whether the job is suspended or denoted as *available*. His/her household expresses the intention to move to another dwelling. The household intends to search for a house with *high intensity.*
- The worker retires and his/her pension is determined. The worker's job is denoted as *available*. His/her household expresses the intention to move to another dwelling. The household intends to search for a house with *low intensity.*

Postcondition: The worker's household denotes whether it intends to search for a new house or not.

C.1.12 Education use case scenario

Use Case Name: Education

Summary: The education career of an individual is determined.

Actor: Individual

Precondition: An individual of age 15 and above.

Description:

The education group to which the individual belongs is determined (none, primary or less, lower secondary, upper secondary, post secondary, tertiary).

Postcondition: The individual is assigned an education career.

C.1.13 Job entry use case scenario

Use Case Name: Job entry

Summary: An individual, who does not already work, finds a job.

Actor: Individual

Precondition: An individual of age 15 who is not a soldier, or an individual who does not work (immigrant or currently unemployed worker).

Description:

1. The individual is simulated to search for a job with *high intensity*.
2. The *Job search* use case is called.
3. If the individual finds a job then his/her salary is determined based on his/her education and type of occupation.
4. The worker's household expresses the intention to move to another dwelling.

Alternatives:

- If the individual does not find a job, then he/she is denoted as unemployed and his/her salary is set to the unemployment allowance. His/her household intends to search for a house with *high intensity.*

Postcondition: The worker's household intends to search for a new house with *low intensity*.

C.1.14 Job change use case scenario

Use Case Name: Job change

Summary: A worker searches for a new job.

Actor: Individual

Precondition: An individual, who already works, is subjected to a probability of searching for a job.

Description:

1. The worker is simulated to search for a job with *low intensity*.
2. The *Job search* use case is called.
3. If the worker finds a new job, then his/her characteristics are updated.
4. The worker's previous job is denoted as *available*
5. The worker's household expresses the intention to move to another dwelling.

Alternatives:

- If the worker does not find a new job then:

3. The worker remains at his/her current job.
4. He/she is subjected to the probability of having a salary rise.
5. If the salary rise takes place then the amount of rise is determined based on the worker's education, type of occupation and previous salary.
6. The worker's household expresses the intention to move to another dwelling.

Postcondition: The worker's household intends to search for a new house with *low intensity*.

C.1.15 Job search use case scenario

Use Case Name: Job search

Summary: An individual, searches for a job.

Actor: Individual

Precondition: An individual is simulated to search for a job.

Description:

1. The individual searches for a job according to the following criteria:
   a. The type of occupation should correspond to the education group to which the individual belongs.
   b. The job offers a higher salary than the one offered by the worker's current job based on the worker's education group and the type of occupation.
   c. The job is located at the city center or at the residence zone of the individual's house or neighboring residence zones.
2. If the individual finds a job that matches his/her criteria, then he/she accepts the job offer.

Alternatives:

- If the individual does not already have a job, then he/she searches for a job according to the **1.a** criteria (the remaining criteria are not taken into consideration).

Postcondition: The individual's job characteristics are updated.

C.1.16 Demand change use case scenario

Use Case Name: Demand change

Summary: The household's housing standards change.

Actor: Household

Precondition: A household that does not have experienced a demographic or income change event is subjected to the probability of searching for a new dwelling due to demand changing.

Description:

- A household whose rent and housing expenditure are lower than the household's housing budget is simulated to request housing in a more attractive neighborhood aiming for socio-economic upgrading.
- A household whose rent and housing expenditure exceed the household's housing budget is simulated to request a less expensive dwelling.

Postcondition: The household intends to search for a new house with *low intensity*.

C.2 The Housing Supply use case scenarios

The Housing Supply sub-model consists of the following use cases:

C.2.1 New construction use case scenario

Use Case Name: New construction

Summary: New dwellings are constructed.

Actor: Dwelling

Precondition: The number of new dwellings to be added to the housing supply set is determined based on the market's mobility and housing shortage.

Description:

1. The location of each new dwelling is determined based on the undeveloped land suitable for development of the market area and the market's housing demand trends.
2. The dwelling's type and services are determined based on the market's housing demand trends.
3. A new dwelling is created and added to the vacated dwellings set, for each newly constructed dwelling.

Postcondition: The house stock is updated.

C.2.2 Structure conversion use case scenario

Use Case Name: Structure conversion

Summary: The structure of a dwelling is converted according to some possible way.

Actor: Dwelling

Precondition: -

Description:

1. If the dwelling is unoccupied and large in size, then the *Split* use case is called.
2. If the dwelling is unoccupied and small in size, then its neighboring dwellings are also examined; if the same conditions apply to one or more neighboring dwellings, then the *Combination* use case is called.
3. The *Renovation* use case is called.

Postcondition: The house stock is updated.

C.2.3 Renovation use case scenario

Use Case Name: Renovation

Summary: A dwelling is renovated.

Actor: Dwelling

Precondition: A dwelling in medium or poor condition is subjected to the probability of being renovated.

Description:

1. The cost and difficulty of the renovation are determined based on the current and targeted construction qualities of the dwelling.
2. If the renovation takes place, then the type of renovation is determined (dwelling maintenance, dwelling extension, number of rooms change etc).
3. The dwelling's characteristics are updated.

Postcondition: The house stock is updated.

C.2.4 Split use case scenario

Use Case Name: Split

Summary: A dwelling is split to two or more smaller dwellings.

Actor: Dwelling

Precondition: A vacant, large dwelling is subjected to the probability of splitting to two or more smaller dwellings.

Description:

1. The cost and difficulty of the split are determined based on how similar the current and targeted structure types and construction qualities are.
2. If the split is profitable[6], then the split takes place.
3. The number of new dwellings is determined based on the size of the original dwelling.
4. The new dwellings are created and added to the vacated dwellings set.

---

[6] The amount that a property owner can invest in a particular year depends on the cash flow the building generates and on current and anticipated levels of neighborhood quality [HUDS].

5. The original dwelling is removed from the dwellings set.

Postcondition: The house stock is updated.

C.2.5 Combination use case scenario

Use Case Name: Combination

Summary: Two or more dwellings are combined to a larger one.

Actor: Dwelling

Precondition: Two or more vacant, small and neighboring dwellings are subjected to the probability of being combined to a larger dwelling.

Description:

1. The cost and difficulty of the combination are determined based on how similar the current and targeted structure types and construction qualities are.
2. If the combination is profitable, then the combination takes place.
3. A new dwelling is created and added to the vacated dwellings set.
4. The original dwellings are removed from the dwellings set.

Postcondition: The house stock is updated.

C.2.6 Demolition use case scenario

Use Case Name: Demolition

Summary: A dwelling is demolished.

Actor: Dwelling

Precondition: A dwelling in poor condition is subjected to the probability of being demolished.

Description:

1. If the dwelling is unoccupied then the demolition takes place.
2. The demolished dwelling is removed from the dwellings set.

Alternatives:

1. If the dwelling is occupied, then the corresponding household expresses the intention to search for a house with *high intensity.* The demolition takes place only after the household moves out

Postcondition: The house stock is updated.

C.2.7 Housing expenditures change use case scenario

Use Case Name: Housing expenditures change

Summary: The housing cost of a dwelling, changes.

Actor: Dwelling

Precondition:

Description:

The housing expenditures (maintenance capital, heat, electricity etc) of the dwelling is determined based on the dwelling's type, size, age and construction quality.

Postcondition: The dwelling's characteristics are updated.

C.3 The Residential Search and Migration use case scenarios

The Residential Search and Migration sub-model consists of the following use cases:

C.3.1 Search and migration use case scenario

Use Case Name: Residential search and migration

Summary: All households whose structure or housing circumstances have changed and all newly formed households as well as the households they stem from, are simulated to search for a new dwelling.

Actor: Household, Dwelling

Precondition: A household whose structure or housing circumstances have changed or a newly formed household.

Description:

1. The *Tenure* use case is called.
2. The *Dwelling Search* use case is called.
3. If the household has moved to a new dwelling then the *Dwelling vacation* use case is called for the dwelling the household occupied (if any).
4. The original household(s) (if any) expresses the intension to move to another dwelling.

Alternatives:

- It is decided the amount of **mortgage**.
- If the household fails to find a dwelling then:
  - If it is an in-migrating household, the household as well as its members is deleted.
  - In any other case, the household is subjected to the probability of *out-migrating*.
    - If the household does not out-migrate then:
      - If it already has a dwelling it remains there.
      - In any other case the household is deleted and its members return to their parental households.

Postcondition: The original household(s) (if any) intends to search for a new dwelling with low search intensity.

C.3.2 Dwelling Search use case scenario

Use Case Name: Dwelling Search

Summary: A household that intends to move to another dwelling searches the list of vacant dwellings.

Actor: Household, Dwelling

Precondition: A household searches for another dwelling.

Description:

1. The household searches the list of vacant dwellings for a dwelling according to the following rules:
   a. Dwellings are defined by seven (7) dimensions: *dwelling price/rent, number of rooms, dwelling type, dwelling age*, *structure condition, amenities (garden, parking etc),* and *neighborhood quality.* The household defines the values of the dimensions the desired dwelling should have, as well as the dimensions that are of importance (primary dimensions) and those that are not (secondary dimensions). The desired dwelling is defined based on the event experienced by the household, the household's income as well as whether the household searches for a dwelling to rent or own.
   b. The *number of dwellings considered* by the household, depends on the *search intensity* of the household. In other words, households with high search intensity consider more dwellings than the ones with low search intensity.
2. If the household decides to accept a vacancy it moves to the new dwelling and the household and dwelling characteristics are updated.

Postcondition: The household denotes whether it has moved to a new dwelling or not.

C.3.3 Dwelling vacation use case scenario

Use Case Name: Dwelling Vacation

Summary: If the household moves to a new dwelling then the dwelling the household occupied is vacated

Actor: Household

Precondition: The household decides to accept a vacancy and moves to the new dwelling

Description:

1. The dwelling is vacated and added to the housing stock.
2. It is decided whether the dwelling will be available for rent or sale.

Postcondition: The house stock is updated.

C.3.4 Housing prices change use case scenario

Use Case Name: Housing Prices Change

Summary: The rent and market value of each dwelling are updated based on housing demand and supply.

Actor: Household, Dwelling

Precondition: -

Description:

- The market value of a dwelling depends on:
  - The region-wide (national-wide) value inflation rate.
  - The municipality's *housing demand* and *vacancy rate* for dwellings having the same dwelling type, size, age, neighborhood quality and construction quality with the dwelling examined.
- In case of a rental dwelling, its rent is determined according to:
  - The region-wide (national-wide) rent inflation rate.
  - The municipality's *housing demand* and *vacancy rate* for renting dwellings that have the same dwelling type, size, age, neighborhood quality and construction quality with the dwelling examined.

Postcondition: The dwellings' characteristics are updated.

Appendix D
The Household and Dwelling Questionnaire

This appendix presents the questionnaire used to collect household and dwelling information for the city of Mytilene, which was in turn used in the development of Mytilene's synthetic population.

HOUSEHOLD – DWELLING QUESTIONNAIRE

A. DWELLING INFORMATION

1. α) AREA …..………………… **b) ADDRESS** ………………………	**2. BUILDING TYPE** ☐ One dwelling house ☐ Two dwelling houses ☐ Three or more dwelling houses	**3. CONSTRUCTION PERIOD** ☐ -1919 ☐ 1919-45 ☐ 1946-60 ☐ 1961-70 ☐ 1971-80	 ☐ 1981-85 ☐ 1986-90 ☐ 1991-95 ☐ 1996-00 ☐ 2000+
4. Renovation year. …………………… **5. Lot Size (m2):** ……	**6. Dwelling Size (m2):…..** **7. Number of bedrooms** ………..	**8. TYPE OF HEAT** ☐ Central ☐ Other ………………… ☐ None	

B. HOUSEHOLD INFORMATION

1. Tenure status: ☐ Owner ☐ Renter ☐ Other …………………..….. Description	**2. Household composition:** Males …………… Females ……………	**3. Household composition January 2004:** Males …………… Females …………… **4. Household composition January 2000:** Males …………… Females ……………

C. HOUSEHOLD HEAD INFORMATION

1. SEX ☐ Male ☐ Female **2. Year of Birth** ………………	**3. Temporarily absent. He/she is:** ☐ Other municipality ………………………….. ☐ Abroad …………………………… Reason ………..……………… …………………………………	**4. Marital Status** ☐ Never Married ☐ Married ☐ Widow ☐ Divorced ☐ Separated ☐ Cohabitate
5. Place of Birth ☐ This municipality ☐ Elsewhere ………………………….	**6. Country Citizenship** ☐ Greece ☐ Greece + Other country ……………………… ☐ Other country ……………….	**7. Municipality Registration** ☐ This municipality ☐ Elsewhere

8. Dwelling of residence in January 2004

☐ Same dwelling

☐ Other dwelling

………………………
Move reason

………………………..
Area

………………………
Municipality/Country

9. Dwelling of residence in January 2000

☐ Same dwelling

☐ Other dwelling

………………………
Move reason

………………………..
Area

………………………
Municipality/Country

10. Dwelling of residence in January 1995

☐ Same dwelling

☐ Other dwelling

………………………
Move reason

………………………..
Area

………………………
Municipality/Country

11. Searching for a new dwelling? ☐ YES ☐ NO

12. If the answer is YES give reason

……………………………………………………………

13. Education Level

☐ PhD	☐ TEI College	☐ Lower Secondary
☐ Masters	☐ Post Secondary	☐ Primary School
☐ University Degree	☐ Upper Secondary	☐ Left Primary School

14. Degree

…………………………
α) University

…………………………
β) School

…………………………
γ) Department

…………………………
δ) Municipality/Country

15. What is your employment status?

☐ Working

☐ Unemployed

☐ Job starter

☐ Schoolchild

☐ Student

☐ Retired

☐ Property incomes

☐ Family Duties

☐ Army service

☐ Other

………………………

16. If the answer in question 15 was not working, which was your employment status the last time you worked?

☐ Working

☐ Unemployed

☐ Job starter

☐ Schoolchild

☐ Student

☐ Retired

☐ Property incomes

☐ Family Duties

☐ Army service

☐ Other…………

<table>
<tr><td>17. What kind of shop, company etc do you work or did you work at?
(Give full description)
……………………………</td><td rowspan="2">19. α) What was your occupation the last time you worked
(e.g. Accountant assistant)
………………………….

β) What are or were the major duties at this work?
(e.g. book keeping)
…………………………</td><td>20. How many hours do you work or did you work per week?
…………………………</td></tr>
<tr><td>18. What was your employment status
☐ Employer
☐ Self-employed
☐ Employee
☐ Family worker</td><td>21. How many people work or worked at the shop, company etc you work or worked?
…………………….</td></tr>
<tr><td rowspan="3">22. Place of work
☐ Not permanent
☐ Permanent
☐ In the house
☐ At the dwelling municipality
☐ At another municipality</td><td rowspan="3">23. What was your employment status last year?
☐ Working
☐ Unemployed
☐ Job starter
☐ Schoolchild
☐ Student
☐ Retired
☐ Property incomes
☐ Family Duties
☐ Army service
☐ Other
………………………</td><td>24. What kind of shop, company etc do you work or did you work at?
……………………………</td></tr>
<tr><td>25. What was your employment status
☐ Employer
☐ Self-employed
☐ Employee
☐ Family worker</td></tr>
<tr><td>26. How many work at the shop, company etc you work or worked?
…………</td></tr>
<tr><td colspan="3">Questions 27 & 28 should be answered only by <u>women</u></td></tr>
<tr><td colspan="2">27. How many children have you given birth to?
Number of children …..
Birth year of the 1st, 2nd & last child …………</td><td>28. Total marriage duration
Years of marriage ……..
Year the first marriage took place ………</td></tr>
</table>

FILL IN THE FOLLOWING INFORMATION FOR <u>EVERY</u> HOUSEHOLD MEMBER

Δ. INFORMATION ….. HOUSEHOLD MEMBER

1. SEX
☐ Male
☐ Female

2 Year of Birth
…………

3. Temporarily absent. He/she is:
☐ Other municipality
……………………..
☐ Abroad
………………………
Reason
……….……………

4. Relation to household head
☐ Husband
☐ Cohabiter
☐ Child
☐ Grandson
☐ Nephew
☐ Father/mother
☐ Father/Mother in Law
☐ Grandfather/ Grandmother
☐ Brother/ Sister
☐ Brother/ Sister in Law
☐ Other ……………

5. Marital Status
☐ Never Married
☐ Married
☐ Widow
☐ Divorced
☐ Separated
☐ Cohabitate

6. Place of Birth
☐ This municipality
☐ Elsewhere
………………………….

7. Country Citizenship
☐ Greece
☐ Greece + Other country
………………………
☐ Other country ………………….

8. Municipality Registration
☐ This municipality
☐ Elsewhere

9. Dwelling of residence in January 2004
☐ Same dwelling
☐ Other dwelling
………………………
Move reason
………………………..
Area
………………………
Municipality/Country

10. Dwelling of residence in January 2000
☐ Same dwelling
☐ Other dwelling
………………………
Move reason
………………………..
Area
………………………
Municipality/Country

11. Dwelling of residence in January 1995
☐ Same dwelling
☐ Other dwelling
…………………
Move reason
…………………
Area
……………………
Municipality/Country

12. Searching for a new dwelling? ☐ YES ☐ NO

13. If the answer is YES give reason ………………………………………………………

14. Education Level ☐ PhD ☐ Masters ☐ University Degree	 ☐ TEI College ☐ Post Secondary ☐ Upper Secondary	 ☐ Lower Secondary ☐ Primary School ☐ Left Primary School
15. Degree ………………………… α) University ………………………… β) School ………………………… γ) Department ………………………… δ) Municipality/Country	**16. What is your employment status?** ☐ Working ☐ Unemployed ☐ Job starter ☐ Schoolchild ☐ Student ☐ Retired ☐ Property incomes ☐ Family Duties ☐ Army service ☐ Other ………………………	**17. If the answer in question 16 was not working, which was your employment status the last time you worked?** ☐ Working ☐ Unemployed ☐ Job starter ☐ Schoolchild ☐ Student ☐ Retired ☐ Property incomes ☐ Family Duties ☐ Army service ☐ Other………………………
18. What kind of shop, company etc do you work or did you work at? (Give full description) ……………………………	**20. α) What was your occupation the last time you worked** (e.g. Accountant assistant) ………………………….	**21. How many hours do you work or did you work per week?** ………………………………………
19 What was your employment status ☐ Employer ☐ Self-employed ☐ Employee ☐ Family worker	**β) What are or were the major duties at this work?** (e.g. book keeping) …………………………	**22. How many people work or worked at the shop, company etc you work or worked?** ……………………………………….
23. Place of work ☐ Not permanent ☐ Permanent ☐ In the house	**24. What was your employment status last year?** ☐ Working	**25. What kind of shop, company etc do you work or did you work at?** ……………………………

<table>
<tr><td>☐ At the dwelling municipality
☐ At another municipality</td><td>☐ Unemployed
☐ Job starter
☐ Schoolchild
☐ Student
☐ Retired
☐ Property incomes
☐ Family Duties
☐ Army service
☐ Other
………………………</td><td>26. . What was your employment status
☐ Employer
☐ Self-employed
☐ Employee
☐ Family worker<hr>27. How many work at the shop, company etc you work or worked? …………</td></tr>
<tr><td colspan="3" align="center">Questions 28 & 29 should be answered only by <u>women</u></td></tr>
<tr><td colspan="2">28. How many children have you given birth to?
Number of children …..
Birth year of the 1st, 2nd & last child …………</td><td>29. Total marriage duration
Years of marriage ……..
Year the first marriage took place ………</td></tr>
</table>