



cuACS

Matching Algorithm Overview



ARCLIGHT ENTERTAINMENT

BRIAN GRICKITES
SHALIN LATHIGRA
BRIAN LEBLANC
AARON RAMOS-LAZETTE

Submitted to:
Dr. Christine Laurendeau
COMP 3004 Object-Oriented Software Engineering
School of Computer Science
Carleton University

MARCH 5, 2019

Table of Contents

1. Introduction.....	2
2. Object Model	3
3. Attributes	4
3.1. Attribute Notes.....	4
3.1.1. Measurements	4
3.1.2. Distance.....	4
3.1.3. Inverse Distance	4
3.1.4. Attribute Categories.....	5
3.2. List of Attributes.....	5
3.2.1. Animal Attributes	5
3.2.2. Client Attributes	6
3.3. Comparing Attributes.....	7
3.4. Summary.....	14
4. Matching Algorithm	16
4.1. Candidate Filter.....	16
4.2. Best-First Search Heuristic	17
4.3. Production System	17
4.4. Pruning System	19
4.4.1. Optional Pruning Performance Increase.....	20
5. Glossary.....	21

Table of Figures

Figure 1 - Object Model	3
Figure 2 - Relationship of Animal to Client attributes	15
Figure 3 - High-level of ACM flow	16
Figure 4 – High-level tree expansion model (Production System)	18
Figure 5 - High-level Pruning System logic flow.....	19

1. Introduction

The Carleton University Animal Care System (cuACS) is responsible for ensuring that sheltered Animals are provided a long-term home that is well suited to them. With that in mind, the Animal Client Matching (ACM) algorithm is responsible for ensuring that Animals are matched with compatible Clients. To do so, the ACM determines a compatibility score for each Animal-Client pair (called a Candidate) and attempts to find Match set (Solution) with the highest mean average for the compatibility scores.

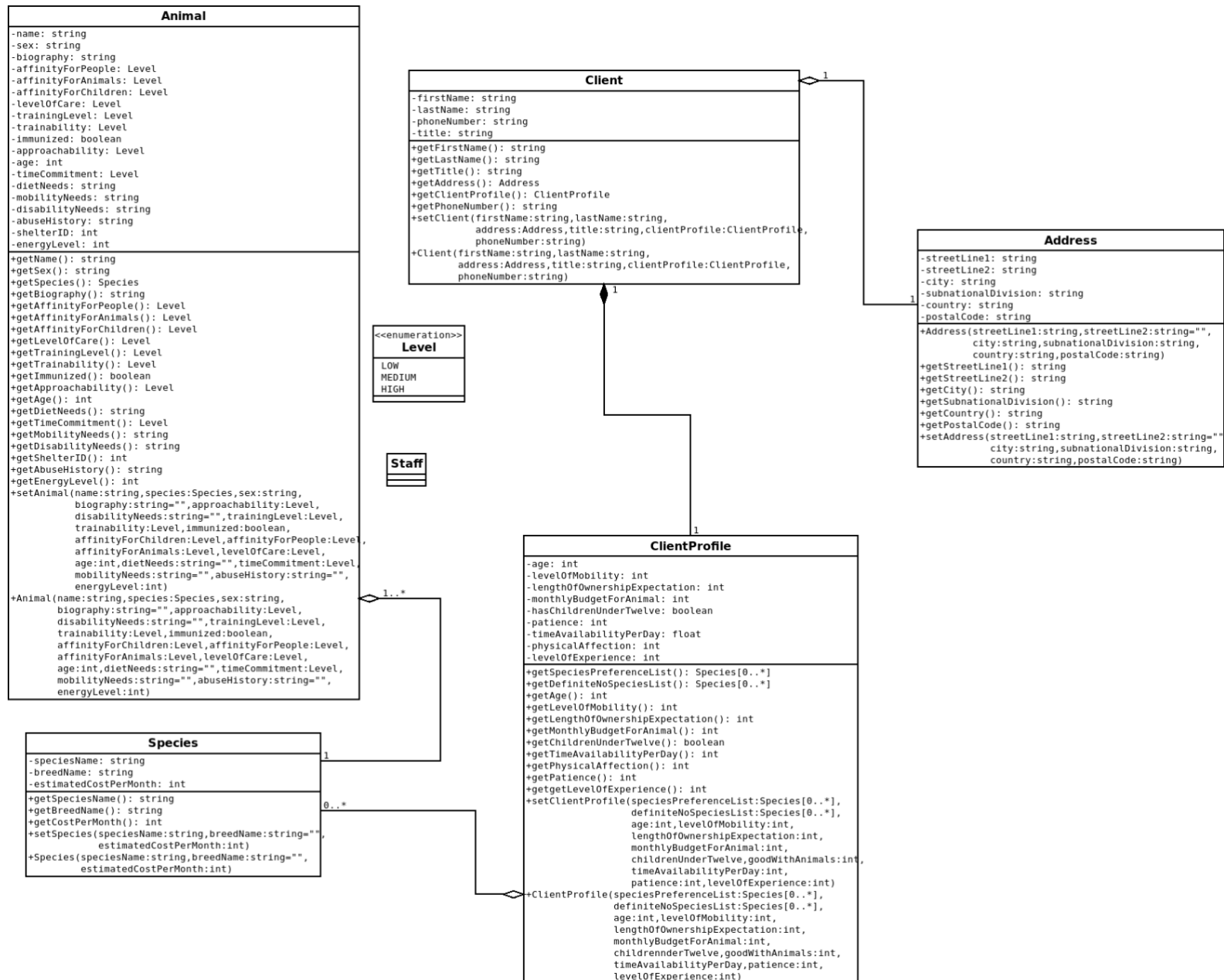
The ACM will attempt to find a suitable Match for all Animals in the shelter and will consider alternate combinations of matches until this requirement is met. However, if the algorithm is forced to choose between matching the Animal with an unsuitable home or keeping the Animal at the shelter until a more compatible home can be found, the algorithm will not choose a home that is not well-suited for the Animal.

This document seeks to outline the utility and constraints of the ACM. Included are summary of changes to the system model, attributes and their inter-relationships that are evaluated by the ACM to determine suitability of Potential Matches. Finally, the finer details of the ACM search algorithm will be explored including filters, heuristics, and other systems used while calculating matches.

2. Object Model

Since the previous iteration of cuACS, some minor adjustments have been made to our system model, illustrated by the diagram below. The primary change involved merging our ClientTraits and Preferences into a single ClientProfile object, broadly defined as containing all Client attributes outside of basic contact info. The new model is as follows:

Figure 1 - Object Model



3. Attributes

3.1. Attribute Notes

The following are a few notes on our groupings and attribute design.

3.1.1. Measurements

An Animal's attributes (measured as 0, 1, 2, which correspond to Low, Medium, High) are measured by the Staff, and as a result will be less accurate than a human Client which can measure their own attributes. Because of this, we have their attributes measured more generally to better account for variability and time of observation by the Staff.

The Clients' attributes (measured as 0, 1, 2, 3, 4) are measured by the Clients themselves and can therefore be allowed to have more specificity than the Animals' measurements.

3.1.2. Distance

- Distance is the measure by which we multiply an Animal's attribute (measured as 0, 1, 2) by two, and subtract from it a human's attribute (measured as 0, 1, 2, 3, 4). The higher the difference (the distance), the worse the Match.
- A distance of 0 indicates that a given Client satisfies all of an Animal's needs with regard to that particular mapping.
- If a Client attribute is higher or equal to two times the Animal's attribute with which it is being mapped, we set the distance as zero, as we are matching specifically with respect to comfort of the Animal.
- Each distance is out of four (this includes inverse distance).

3.1.3. Inverse Distance

We sometimes use the inverse of an Animal's attribute ($4 - \text{The Attribute} \times 2$) in order to better represent the distance. For example, when comparing a Client's Physical Affection with an Animal's Affinity for People, we want to subtract Patience from how much the Animal *dislikes* people. We estimate this by taking the Animal's Affinity for People and subtracting two times that value from four. Inverse distance always refers to the inverse of the Animal's attribute being mapped.

3.1.4. Attribute Categories

There are three groups: Social Needs, Physical Needs, and Other Needs. We assign each mapping a value of Low, Medium, and High, depending on the weight we will assign to that relation for the final value.

- **Physical Needs:** Needs relating to physical care of the Animal, such as walking them, cleaning litter boxes, training and so on.
- **Social Needs:** Needs relating to intangible social aspects of the Animal being mapped to aspects of the Client. Generally related to the Animal's personality.
- **Other Needs:** Needs specifically considering environmental characteristics of the Client's home, such as budget, children being present in the home, as also accounts for life expectancy.

3.2. List of Attributes

The following attributes are used by the ACM to determine the compatibility of any given Animal and Client.

3.2.1. Animal Attributes

- **Species Name:** The given name of a species.
- **Affinity for Children:** A rating denoting how well behaved with children an Animal is. Encapsulates gentleness with children who may be overzealous regarding their physical affection towards the Animal.
- **Affinity for People:** A rating denoting how receptive an Animal is to any form of interaction with people.
- **Affinity for Animals:** A rating denoting how easily an Animal remains in proximity with other non-human Animals which share the same space.
- **Level of Care:** Encapsulates any physical needs required by the Animal, including feeding, exercise, medical needs, etc.
- **Training Level:** A rating of how trained an Animal is with regards to living in a household environment (for example, house training regarding keeping off couches). An Animal that does not require training is considered to be fully trained (for example, a fish).

- **Trainability:** Refers to the Animal's willingness and ability to be trained out of bad behaviours and respond to human command. An Animal that cannot be trained is considered to be fully trainable (for example, a fish).
- **Approachability:** Refers to an Animal's willingness to be approached by a human.
- **Time Commitment:** The estimated time commitment (by level) that an Animal requires daily.
- **Energy Level:** Approximates the Animal's energy level with regard to physical activity.
- **Estimated Cost Per Month:** The estimated cost per month which an Animal requires in terms of monetary expenditure.
- **Life Expectancy:** Estimated by subtracting Animal age from the estimated length of life for that species. The expected years which an Animal has left to live.

3.2.2. Client Attributes

- **Patience:** The level concerning the patience of a Client in dealing with an Animal (for example when training).
- **Physical Affection:** The level of physical affection a Client would prefer to display towards their Animal.
- **Previous Experience:** The level of previous experience in pet ownership which a Client has.
- **Level of Mobility:** The level of mobility which the Client possesses. Includes how sedentary they may be, as well as any physical mobility restrictions they may have.
- **Time Availability per Day:** The time (in minutes) which a Client is available to focus exclusively on the Animal in question per day.
- **Species Preference List:** A list of species which the Client would prefer to be matched with.
- **Definite No List:** A list of species that the Client is voluntarily excluding themselves from being matched with.
- **Has Pets:** A Boolean value denoting whether or not a Client currently has pets.

- **Has Children Under Twelve:** A Boolean concerning whether or not the Client has children under twelve years of age.
- **Estimated budget:** The Client's monthly budget which they are willing to spend on their Animal.
- **Length of ownership Expectation:** The length in years which a Client is expecting to own the Animal they adopt.

3.3. Comparing Attributes

The following represents the model by which Animal and Client attributes are compared with one another to determine compatibility.

Animal Species Name -> Client Definite No List

Weight: Immediate Unsuitability Test

Group: N/A

- The Client generates their own 'definite no' list. This can be based on anything from allergies (for example to dogs), or simply to personal reasons (perhaps they are terrified of dogs). In either case, this will cause the comparison to return a value of total unsuitability (-1).

Animal Affinity for Children -> Client Children Under Twelve

Weight: High (Immediate Unsuitability Test)

Group: Environment Constraints

- An Animal with a 0 Affinity for Children would not do well in a house with children and could potentially be hostile to them. As this is the case, should an Animal have an Affinity for Children of 0, and a house has children, the comparison will return a value of total unsuitability (-1).
- Otherwise, we take the inverse of Affinity for Children ($4 - \text{Affinity for Children} \times 2$) and add that directly to personality if the house possesses children.
- If there are no children in the house, the distance is set as zero.
- We weight this as high, as we assume that any given Client will prioritize the wellbeing and comfort of their children over any adopted Animal, making the Animal's suitability with their children of high priority.

Animal Affinity for People -> Client Patience:

Weight: Medium

Group: Social Needs

- A Client with high Patience could cope with an Animal not particularly enjoying people, by (for example) slowly earning their trust. On the other hand, a person with low Patience would not have to wait to create a bond with an Animal which has high Affinity for People. As a result, we take the inverse distance of Patience and Affinity for People.
- We weight this as medium as while a patient Client may be frustrated with an Animal's disinclination towards people, their common sense would prevent them from harassing or irritating the Animal.

Animal Affinity for People -> Client Physical Affection:

Weight: High

Group: Social Needs

- An Animal with high Affinity for People would pair best with a Client with high Physical Affection, and vice versa. For example, a cat that does not like to be pet, would pair poorly with a Client who enjoys petting their Animal. As a result of this, we use the inverse distance between Affinity for People and Physical Affection.
- We weight this as high because an Animal which dislikes physical affection paired with an owner who prefers this would be a great disservice to both, as people who value physical affection tend to do so highly when it comes to Animals.

Animal Level of Care -> Client Previous Experience:

Weight: Medium

Group: Social Needs

- An Animal requiring a high Level of Care would pair the best with a Client who has Previous Experience owning Animals. Whether it is medical conditions, or simply an Animal which requires a lot of attention, Previous Experience would help a given Client make proper decisions concerning their Animal. For this mapping we use the distance between Previous Experience and Level of Care.
- We weight this as medium as while previous experience can inform some decisions, common sense and Google can make up for lack of experience rather effectively.

Animal Trainability -> Client Patience:

Weight: Medium

Group: Social Needs

- An Animal with low Trainability would be best served with a Client with high Patience. If, for example, an Animal has difficulty learning when and where to defecate, a Client with low Patience could quickly get frustrated with this. As a result, we use inverse distance for evaluating this mapping.
- We weight this as medium as a patient Client may be able to take the time to slowly train a pet who is comparatively untrainable, but they may not know or understand the best ways to train any given Animal, despite having the patience to do so.

Animal Trainability -> Client Previous Experience:

Weight: Medium

Group: Social Needs

- An Animal with low Trainability would be best served with a Client with high Previous Experience. If, for example, an Animal has difficulty learning when and where to defecate, a Client with low Previous Experience may not know how to effectively teach them to use the litter box. As a result, we use inverse distance for evaluating this mapping.
- We weight this as medium as an experienced Client may be able to know or understand the best ways to train any given Animal, they may not have other intangible assets (such as patience) to do so.

Animal Approachability -> Client Patience:

Weight: High

Group: Social Needs

- An Animal with low Approachability would be served best by a Client with high Patience. If, for example, a cat is skittish upon people approaching it, an owner with low Patience would perhaps chase and distress the cat, whereas one with high Patience would slowly work on having the cat be comfortable around them. As a result, we use the inverse distance for this mapping.
- We weight this as high, as a highly skittish or aggressive Animal could require an immensely patient owner to eventually win their trust and be comfortable with them. An impatient owner could cause stress in an Animal with their repeated attempts at getting the Animal to like them.

Animal Approachability -> Client Previous Experience:

Weight: Low

Group: Social Needs

- An Animal with low Approachability would be served best by a Client with high Previous Experience. If, for example, a cat is skittish upon people approaching it, an owner with low Previous Experience would perhaps be unaware of how to best make the cat comfortable with them, whereas one with high Previous Experience would slowly work on having the cat be comfortable around them using techniques they may be aware of or have used in the past. As a result, we use the inverse distance for this mapping.
- We weight this as low as despite a Client having their approach of a given Animal informed by their experience, a Client may quickly intuit the Animal's preferred manner of approach even if they have no previous experience.

Animal Species Name -> Client Species Preference List:

Weight: N/A

Group: Social Needs

- If an Animal's Species is present in a Client's Preferred Species List, this simply includes a flat 'bonus' in the Personality group ($-4 * 0.1$) which reduces the overall distance.

Animal Training Level -> Client Previous Experience:

Weight: High

Group: Physical Needs

- An Animal with a low Training Level would best be served with a Client which has Previous Experience with Animals, for the purposes of handling an untrained Animal and (potentially) training it. As a result, we use inverse distance for this mapping.
- We weight this as high as an owner who has previous experience with Animals should be able to draw on that experience to better train and control an Animal with minimal training. For example, an owner which has trained numerous dogs to keep off of a couch, should be able to do so more easily than one who had not.

Animal Level of Care -> Client Level of Mobility

Weight: High

Group: Physical Needs

- As Level of Care includes required physical activity, the Client's Level of Mobility is necessary to know. A dog which requires long walks would be, for example, would be poorly paired with a mostly sedentary individual. As a result, we measure this mapping using distance.

- We weight this as high as an owner to an Animal which requires significant physical care should have the mobility to allow them to serve the Animal's needs. A low level of mobility paired with high level of care may lead to an Animal not being fulfilled in terms of physical activity, or sanitary needs (such as litter boxes) which would be very bad for any creature.

Animal Training Level -> Client Level of Mobility:

Weight: Low

Group: Physical Needs

- An Animal which requires significant training would be poorly paired with an individual who has low levels of mobility, either by choice or because of other circumstances. This is because to train an Animal requires potentially significant movement on the Client's part. We measure this using inverse distance.
- We weight this as low as training level does not inform trainability, and an Animal which is untrained may be easily trained by a Client even with limited mobility.

Animal Training Level -> Client Patience:

Weight: High

Group: Physical Needs

- An Animal which requires significant training would be poorly paired with an individual who has low Patience. This is because to train an Animal requires potentially significant patience on the Client's part if an Animal has minimal training, as training can be tedious and repetitive. We measure this using inverse distance.
- We weight this as high because an untrained Animal can be potentially frustrating for a Client with low patience overall. While training level does not inform trainability, simply dealing with an Animal which is not familiar with living in a home environment can be trying.

Animal Training Level -> Client Time Availability per Day:

Weight: Medium

Group: Physical Needs

- An Animal which requires significant training would be poorly paired with an individual who has low Time Availability Per Day. This is because to train an Animal requires potentially significant time on the Client's part, if an Animal has minimal training, as training can be time consuming and slow. We measure this using inverse of the Animal training level and subtracting from that the Client's Time Availability Per Day divided by 30.

- We weight this as medium as while important, a Client could potentially train an Animal in a smaller amount of time depending on the Animal's trainability.

Animal Time Commitment -> Client Time Availability Per Day:

Weight: High

Group: Physical Needs

- Simply, an Animal with high Time Commitment would be ill suited to a Client with low Time Availability Per Day.
- We measure this by setting each level for the Time Commitment attribute to a time in minutes and determining from that if the Client's Time Availability Per Day can satisfy the Animal's demand for time informed by the Time Commitment. If it cannot, we subtract from the value informed by Time Commitment the amount of time the Client is available, and we use that to compute a percentage of time satisfied which is then multiplied by four.
- We weight this as high, as an Animal's time commitment should be fulfilled if at all possible, to best maintain its quality of life from an exercise and care perspective.

Animal Trainability -> Client Time Availability Per Day:

Weight: Medium

Group: Physical Needs

- An Animal which has high Trainability would be more easily trained by an individual who has low Time Availability Per Day. This is because to an Animal with a high degree of Trainability would more easily take to the commands and instruction of the given Client. This distance is calculated by taking the Animal's inverse Trainability, and then subtracting from that the Client's Time Availability Per Day divided by thirty.
- We weight this as medium as, while important, the Animal may be previously trained to some degree.

Animal Energy Level -> Level of Mobility:

Weight: High

Group: Physical Needs

- An Animal with a high energy level will require an owner which is equally energetic, in order to keep it in shape, and keep up with it. We measure this via the distance between Energy Level and Level of Mobility.

- We weight this as high as physical activity for most Animals is incredibly important, and to be able to satisfy an Animal's desire for physical activity is a key criterion for owning any Animal.
- We weight this as high as an Animal's physical needs regarding exercise are paramount to most Animal's wellbeing, and a Client with limited mobility may not be able to satisfy this for the Animal.

Animal Level of Care -> Time Availability Per Day:

Weight: High

Group: Physical Needs

- An Animal requiring a high Level of Care necessary would be ill matched to a Client with low Time Availability Per Day, as caring for an Animal takes, potentially, significant time.
- We measure this by setting each level for the Level of Care attribute to a time in minutes and determining from that if the Client's Time Availability Per Day can satisfy the Animal's demand for time informed by the Level of Care. If it cannot, we subtract from the value informed by Level of care the amount of time the Client is available, and we use that to compute a percentage of time satisfied which is then multiplied by four.
- We weight this as high because the higher the level of care, naturally, the higher the time commitment would have to be from the Client. A Client who cannot fulfill the time requirements from any Animal would significantly negatively impact the Animal's quality of life.

Animal Estimated Cost Per Month -> Client Estimated Budget:

Weight: High

Group: Environment Constraints

- An appropriate home for an Animal must be able to supply it with all necessary food, toys, medical care, and other monetary considerations not mentioned here. If an Animal costs more than a home is potentially willing to spend, this could indicate a poor Match.
- The distance here is calculated by calculating a percentage of how much money the Animal requires which is unsupplied by the Client, and assuming a maximum of 25%, calculates a value by multiplying that percentage by sixteen. If it is above the 25% threshold, the Match is deemed wholly unsuitable.
- We weight this as high because a Client who cannot support an Animal monetarily would almost definitely be a poor Match. That being said, there is the possibility of the Client penny pinching, and making the Match work regardless which is why we

determined a threshold which the Match, while not ideal, would not be a full deal breaker.

Animal Affinity for Animals -> Client Has Other Pets:

Weight: High

Group: Environment Constraints

- If a Client has other pets, an Animal with low Affinity for Animals would be a poor choice for the home. Therefore, this distance is calculated as the inverse of the Animal's Affinity for Animals, only if the home does in fact contain other Animals.
- We weight this as high because, while an Animal can be segregated from other Animals, if other Animals are present in the home it would be anything but ideal for an Animal which dislikes other Animals.

Animal Species Length of Life Estimate – Animal Age -> Client Length of Ownership

Expectation:

Weight: High

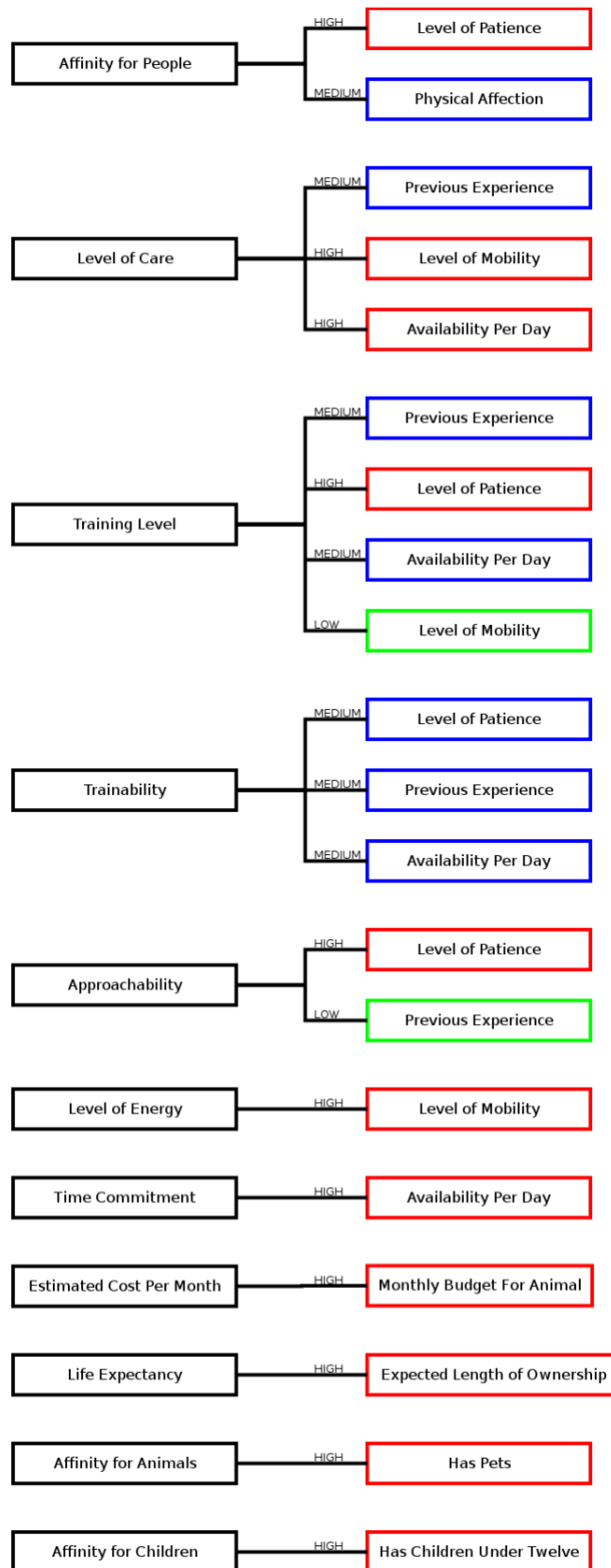
Group: Environment Constraints

- A family looking for a dog to live ten plus years would be ill suited to a dog in the twilight of its life. This distance is calculated by taking the difference between the Animal's estimated years left with the Client's Length of Ownership Expectation. If this value is greater than four, the Match is deemed unsuitable.
- We weight this as high because an Animal in their twilight years would not do well to be adopted by a family looking for a lifetime companion. In the reverse, an Animal with a long life ahead of it may not do well to be adopted by a Client who is looking for a companion for a year or two because of whatever circumstances inform that decision. The latter occasion could in fact incline the Client to return their Animal to some shelter.

3.4. Summary

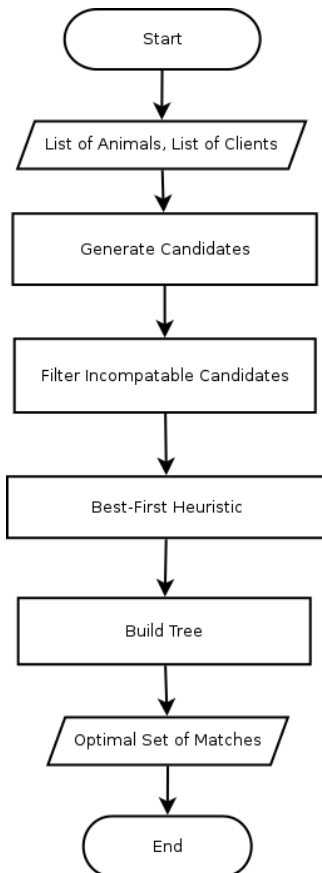
The diagram on the following page seeks to summarize the relationships between Animal attributes (pictured left) and Client attributes (pictured right). Colour is used to denote the weight of the relationship, with green being Low, blue being Medium, and red being High.

Figure 2 - Relationship of Animal to Client attributes



4. Matching Algorithm

Figure 3 - High-level of ACM flow



The ACM, as a whole can be summarized by the following high-level model depicted left.

The ACM seeks to provide an optimal set of matches by using the compatibility factor for Candidates to find the greatest net sum over all of them. There are, however, a few exceptions to this rule that will be elaborated on momentarily. After using the above techniques to determine the compatibility factors of each Candidate, they are manipulated by the following four processes to determine the greatest set of matches: Candidate Filter, Heuristics, Production System, and Pruning System.

4.1. Candidate Filter

The Candidate Filter seeks to discard unacceptable matches prior to the Production System performing the bulk of the work. In doing so, it reduces the number of combinations the system has to account for and, resultantly, the size of the graph reduces as well.

The filter takes the Candidates and a minimal threshold, say 30% for the purpose of this example. Any Candidate found to have a compatibility score lower than this threshold of 30% will be immediately discarded and not considered throughout the rest of the algorithm. This helps to set a hard standard for matches, with the ultimate purpose of preventing Solutions that have a high compatibility average but contain many lackluster matches. This threshold may be adjusted as needed, with higher thresholds yielding Solutions with lower statistical variance but increasing the chance that some Animals or Clients may be completely excluded from the Solution because their compatibility level was not tolerated by the filter.

4.2. Best-First Search Heuristic

This Heuristic is used to prioritize which Potential Solutions get evaluated first. It assumes that Potential Solutions that contain the most exceptionally compatible Candidates are more likely to yield a better set of matches. For this reason, these Potential Solution paths will be explored first while less desirable matches will be explored later on. This prioritization is what is known as a Best-First Search.

From here, the Production System handles the exploration of Potential Solutions. This search, assuming the optional Pruning System threshold is not enabled, will yield the best possible Solution, as determined by the mean average across the Matches.

4.3. Production System

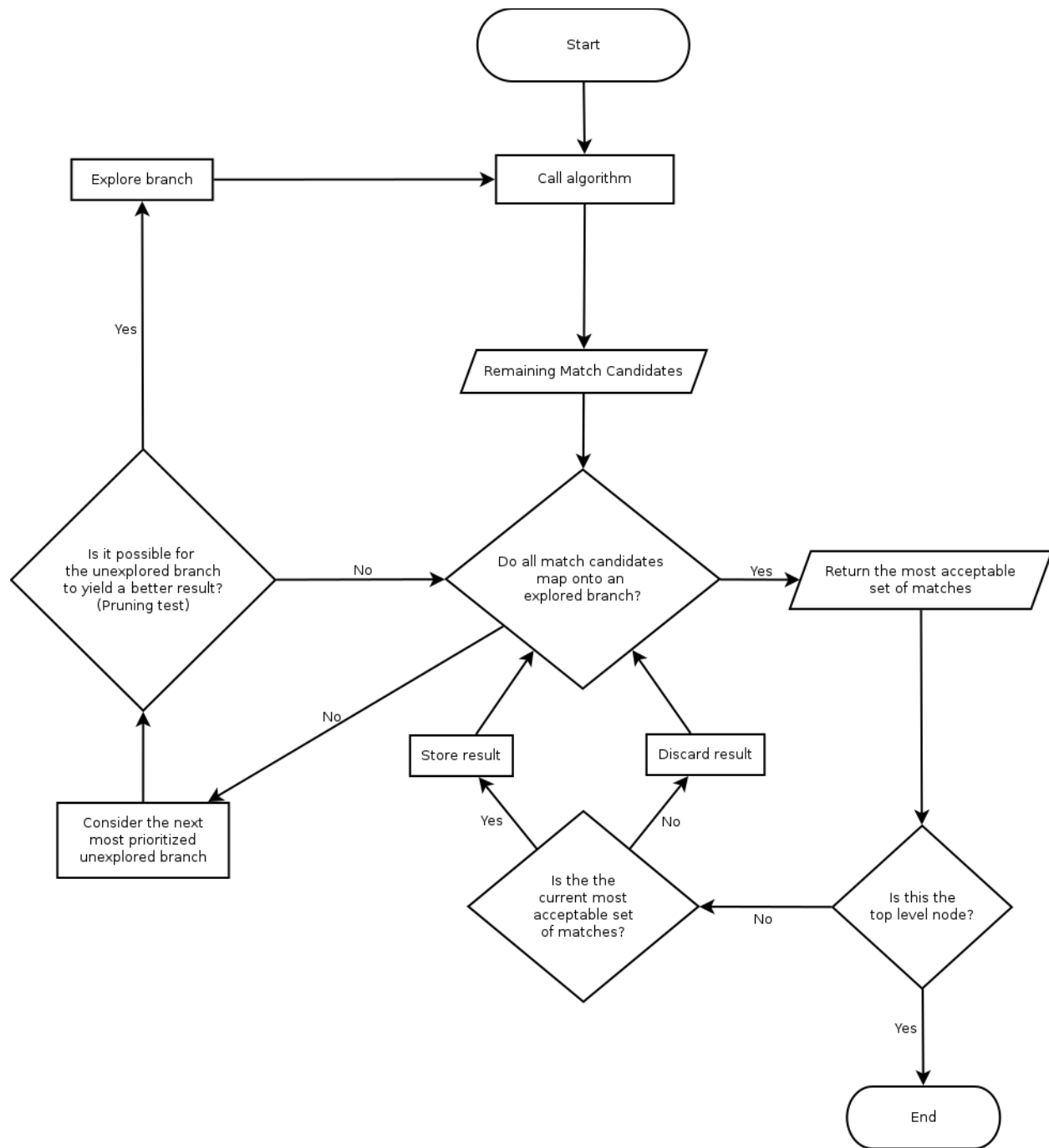
The Production System reflects the present state of the search and outlines the rules to proceed to the next generation(s) of the system, where each new generation is represented as a node further down on the problem tree. Each node reflects a unique state of the search tree and each new generation a reduction in the problem size. In other words, each generation further approaches a terminating state (base case).

At this point in the algorithm, any subset of the Candidates that do not contain the same Animal more than once or the same Client more than once are considered to be Potential Solutions.

Using the Best-First-Search Heuristic describes above, the Production System will attempt to evaluate any Potential Solution that has not previously been considered along the search path. The Production System will, upon selecting the next most promising Candidate, pass on Candidates to a new generation of the Production System to evaluate only if they compare neither the selected Candidate's Animal nor Client.

The following is a high-level representation of the tree building model:

Figure 4 - High-level tree expansion model (Production System)

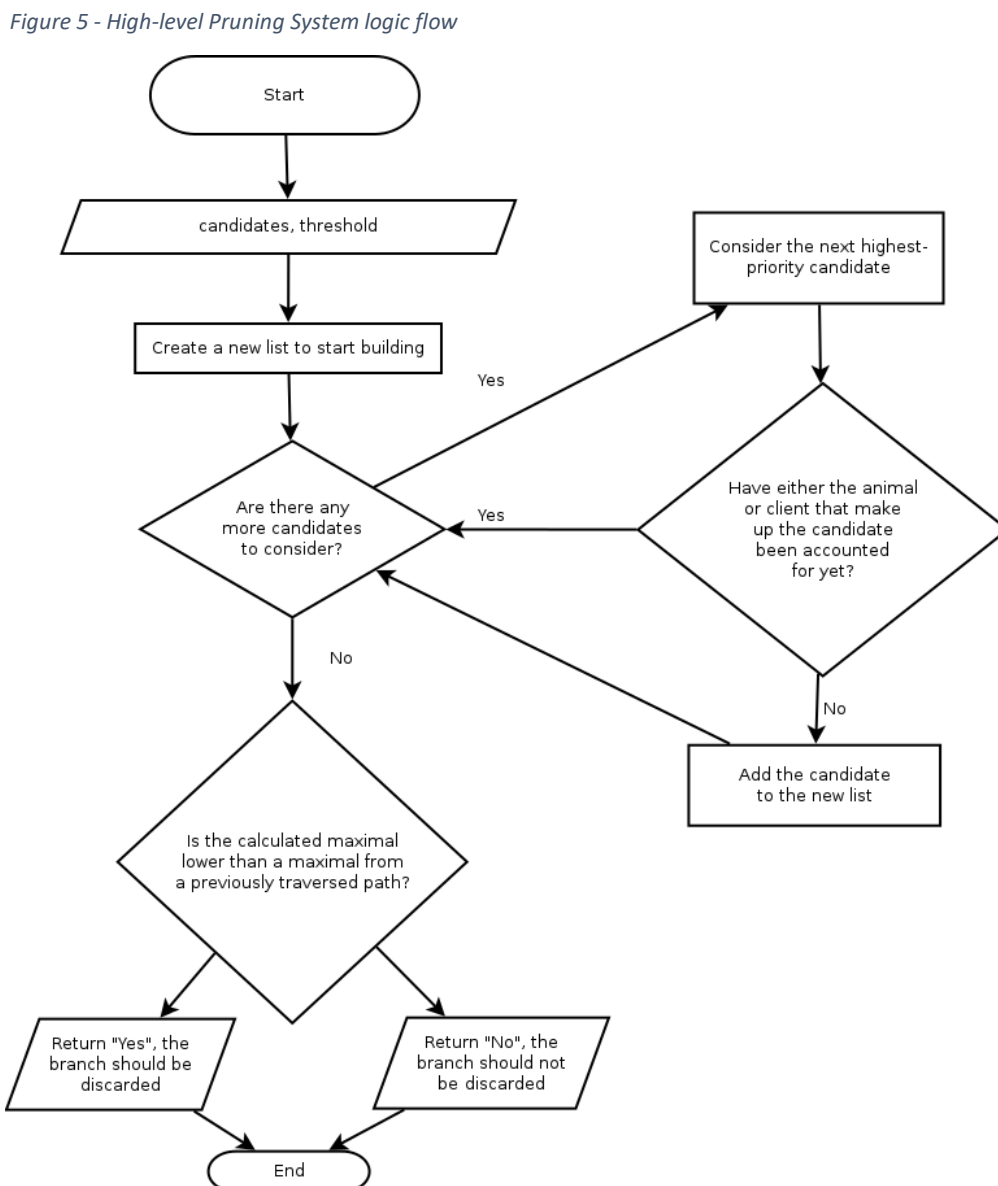


4.4. Pruning System

While the Heuristic works to prioritize the evaluation of Potential Solutions that are assumed to be preferable, it does not in isolation reduce the size of the overall problem. This is where the Pruning System comes in. Any given node in the search tree keeps track of the most exceptional Solution given the remaining unexplored nodes.

Given that current optimal sub path the Pruning System considers whether the next unexplored sub path could, given the best-case scenario, provide a more optimal path than that already discovered. If not, then the Pruning System will discard this path before the Production System even has a chance to evaluate it. Using this technique allows the system to reduce the problem size.

The following diagram models the logic behind the Pruning System:



4.4.1. Optional Pruning Performance Increase

Working under the worst-case scenario that the Pruning System is still not enough to improve the performance of the ACM, an additional optional, modifiable threshold will define an acceptable minimum for the algorithm to produce. Similar to but distinct from the minimum threshold defined in the Candidate Filter, this threshold is used to determine the average acceptable minimum of the entire Solution. The Pruning System would discard options that would not meet this threshold.

If this option is enabled, the algorithm will still look to find a quality set of matches but will terminate once a Solution has been found that meets the minimum requirement determined by the threshold, rather than finding the best possible Solution. This is a trade-off that should only be considered if a shelter housing many Animals and/or having many Clients seeking to adopt would cause too many comparisons to be made.

5. Glossary

Provides a description of key terms used throughout the document.

Staff (Member)	Shelter user with access to the Management account.
Client	Shelter user given a Client profile page by a Staff Member.
Animal	Shelter Animal given an Animal profile page by a Staff Member.
Animal-Client Match (“Match”)	A pairing of a compatible Animal and Client suggested by the ACM Algorithm. A member of the Solution Set.
Candidate	A potential but unconfirmed Match marked for consideration prior to the execution of the ACM Algorithm.
ACM Algorithm	Animal Client Matching Algorithm. Suggests a Match Solution. This optimization for this Solution is determined by a combination of the Search Heuristic, the Production System and the Pruning System.
Solution (Set)	An optimized list of compatible Matches generated by the ACM.
Potential Solution	A set of matches that the system has not yet determined to be optimal.
Production System	How the algorithm determines what subtrees to evaluate.
(Search) Heuristic	A strategy for arriving at the Solution quicker.
Pruning System	The system which decides whether a particular search path can be ignored. Ignoring a path guarantees that it could not have yielded a Solution better than any previous Solutions found.