```
                                        good += x;


        IF exact != NULL THEN return(exact);
        IF good != NULL THEN return(good);
        return(userQuery(dlist));                                        170
}


userQuery(dlist SET OF DN): SET OF DN
{
        -- pass back up for manual checking
        -- user can strip all matches to force progres....
}

                                                                         180
head()    -- return first element of list
tail()    -- return list with first element removed
length()  -- return size of list
last()    -- return last element of list
```

Figure 3: Matching Algorithm

```
                return(search(base, FALSE, s, {OrgUnitName},
                        {substring, approx}, {OrgUnit}));
        ELSE IF present(base, OrganisationName) THEN                    110
                return(search(base, FALSE, {OrgUnitName,
                        LocalityName}, {substring, approx},
                        {Organization, OrgUnit, Locality}));
        ELSE IF present(base, LocalityName) THEN
                return(search(base, FALSE, s, {OrganisationName},
                        {substring, approx}, {Locality});
        ELSE
                return(search(base, false, s, {OrganisationName,
                        LocalityName}, {substring, approx},
                        {Organisation, Locality}));                     120
}


present(d: DN; t: AttributeType): BOOLEAN
{
        FOR x IN d DO
                IF x.type == t THEN return(TRUE);
        return(FALSE);
}
                                                                        130
leafSearch(base: DN; s: String; subtree: BOOLEAN)
{
        return(search(base, subtree, s, {CommonName, Surname,
                UserId}, {substring, approx}));
}


search(base: DN; subtrees: BOOLEAN; s: string;
        alist SET OF AttributeType; matchtypes SET OF MatchType
        objectClasses SET OF ObjectClass OPTIONAL): SET OF DN           140
{
        -- mapped onto Directory Search, with OR conjunction
        -- of filter items

        return dNSelect (s, search-results, alist);
}

read(base: DN; alist SET OF AttributeType): SET OF Attribute;
{
        -- mapped onto Directory Read                                   150
        -- Types repeated to deal with multiple values
        -- This would be implemented by returning selected info
        -- with the search operation
}

dNSelect(s: String; dlist SET OF DN; alist: SET OF AttributeType): SET OF DN
{
        exact, good: SET OF DN;

        FOR x IN dlist DO                                               160
                IF last(DN).Value == s THEN
                        exact += x;
                ELSE IF FOR y IN read(x, alist) DO
                        IF y.value == s THEN
```

```
        IF matches != NULL THEN
                return(matches);
        ELSE
                return(envMatch(p, tail(e));
}


purportedMatch(base: DN; p: PurportedName): SET OF DN
{
        s: String = head(p);                                            60
        matches: SET OF DN = NULL;

        IF length(p) == 1 THEN
                IF length(base) == 0 THEN
                        IF (matches = rootSearch(s)) != NULL THEN
                                return(matches);
                        ELSE return(leafSearch(base, s, FALSE);
                ELSE IF length(base) == 1 THEN
                        IF (matches = intSearch(base, s)) != NULL THEN
                                return(matches);                        70
                        ELSE return(leafSearch(base, s, FALSE);
                ELSE
                        IF (matches = leafSearch(base, s, TRUE)) != NULL THEN
                                return(matches);
                        ELSE return(intsearch(base, s);


        IF length(base) == 0 THEN
                FOR x IN rootSearch(s) DO
                        matches += (purportedMatch(x, tail(p));         80
        ELSE
                FOR x IN intSearch(base, s) DO
                        matches += (purportedMatch(x, tail(p));
        return(matches);
}


-- General.  Might need to tighten the filter for short strings,
-- in order to stop being flooded.  Alternatively, this could be
-- done if the loose search hists a size limit                          90

rootSearch(s: String): SET OF DN
{
        IF length == 2 THEN
                return(search(NULL, FALSE, s, {CountryName,
                        FriendlyCountryName, OrganizationName},
                        {exact}, {Country, Organisation}));
        ELSE
                return(search(NULL, FALSE, s, {OrganizationName,
                        FriendlyCountryName}, {substring, approx},      100
                        {Country, Organisation}));
}


intSearch( base: DN; s: String)
{
        IF present(base, OrgUnitName) THEN
```

A Pseudo-code for the matching algorithm

The following pseudo-code is intended to clarify the matching algorithm. The language uses ASN.1 data types, with flow control 'C'-like, but with keywords upper–cased.

```
PurportedName ::= SEQUENCE OF String
                    -- simplication, as attribute types can optionally be
                    -- specified

                    -- Each element of the Purported Name is a string
                    -- which has been parsed from the BNF

Attribute ::= SEQUENCE {                                                    10
          type OBJECT IDENTIFIER,
          value ANY }

RDN ::= Attribute -- simplification, as can be multi-value

DN ::= SEQUENCE OF RDN

Environment ::= SEQUENCE OF DN

                                                                            20
EnvironmentList ::= SEQUENCE OF SEQUENCE {
                              lower-bound INTEGER,
                              upper-bound INTEGER,
                              environment Environment }


friendlyMatch(p: PurportedName; el: EnvironmentList):  SET OF DN
{
                                        -- Find correct environment
                                                                            30
          IF length(el) == 0 THEN return(NULL);

          IF length(p) <= head(el).upper-bound
                              && length(p) >= head(el).lower-bound THEN
                    return envMatch (p, head(el).environment);
          ELSE
                    return(friendlyMatch(p, tail(el));
}

                                                                            40
envMatch(p: PurportedName; e: Environment):  SET OF DN
{
                                        -- Check elements of environment
                                        -- in the defined order

          matches: SET OF DN;

          IF length(e) == 0 THEN return(NULL);

          matches = purportedMatch(head(el).DN, p)                          50
```

[Neu89] G.W. Neufeld. Descriptive names in X.500. In *SIGCOMM 89 Symposiun Communications Architectures and Protocols*, pages 64–71, September 1989.

[Pet88] L.L. Petersen. The profile naming service. *ACM Transactions on Computing Systems*, 6(4):341–364, November 1988.

Untyped and Ordered This is the type of name proposed here (with some extensions to allow optional typing). It is seen as meeting the key user requirement of disliking typed names, and is efficient to implement.

Typed and Unordered This sort of name is proposed by others as the key basis for user friendly naming. Neufeld shows how X.500 can be used to provide this [Neu89], and Peterson proposes the Profile system to provide this [Pet88]. The author contends that whilst typed naming is interesting for some types of searching (e.g., yellow page searching), it is less desirable for naming objects. This is born out by operational experience with OSI Directories [Kil88].

Untyped and Unordered Surprisingly this form of name can be supported quite easily. However, a considerably gain in efficiency can be achieved by requiring ordering. In practice, users can supply this easily. Therefore, this type of name is not proposed.

10 Proposal

This approach should be tested out experimentally, with the QUIPU implementation. This will involve:

- An interface to generate strings in this form from DNs, with control over
  - Folding (line, column, or max width)
  - Presence of commas at EOL
  - A DN to be abbreviated against
- A routine to resolve a purported name in the context of an environment.

The Distribution List interface seems to be an ideal system to experiment with.

A specification should then be progressed, and made widely available. Perhaps as a RARE Document or RFC.

References

[CCI88] The directory - overview of concepts, models and services, December 1988. CCITT X.500 Series Recommendations.

[Kil88] S.E. Kille. The THORN large scale pilot exercise. *Computer Networks and ISDN Systems*, 16(1):143–145, September 1988.

[Kil89] S.E. Kille. New name forms, May 1989. ISO/IEC/JTC 21/ WG4/N797 UK National Body Contribution to the Oslo Directory Meeting.

```
-> t hales, csiro, australia
Found good match(es) for 'australia'
Found exact match(es) for 'csiro'
Please select from the following:
   Trevor Hales, OC, HPCC, DIT, IICT, CSIRO, AU [y/n] ? y
The following were matched...
   Trevor Hales, OC, HPCC, DIT, IICT, CSIRO, AU

-> g michaelson, queensland, au
Found exact match(es) for 'au'
Please select from the following:
   University of Queensland, AU [y/n] ? y
   Axolotl, AU [y/n] ? n
Please select from the following:
   George Michaelson, Prentice Computer Centre, University of Queensland, AU
[y/n] ? y
   Manager, University of Queensland, AU [y/n] ? n
The following were matched...
   George Michaelson, Prentice Computer Centre, University of Queensland, AU

-> r needham, cambridge
Found good match(es) for 'cambridge'
Please select from the following:
   Roger Needham, Computer Lab, Cambridge University [y/n] ? y
The following were matched...
   Roger Needham, Computer Lab, Cambridge University

-> kirstein
Found good match(es) for 'kirstein'
The following were matched...
   Peter Kirstein
```

Figure 2: Example usage of User Friendly Naming

7 Support of OSI Services

The major focus of this work has been to provide a mechanism for identifying Organisations and Users. A related function is to identify applications. Where the Application is identified by an AET (Application Entity Title) with an RDN of Common Name, this proposal leads to a natural usage. For example, if a filestore in named "gannet", then this could easily be identified by the name:

```
Gannet, Computer Laboratory, Cambridge University, GB
```

In normal usage, this might lead to access (using a purported name) of:

```
FTAM gannet,cambridge
```

A second type of access is where the user identifies an Organisation (Organisational Unit), and expects to obtain a default service. The service is implied by the application, and should not require any additional naming as far as the user is concerned. It is proposed that this is supported by User Friendly Naming in the following way.

1. Determine that the purported name identifies a non-leaf object, which is of object class Organisation or Organistaional Unit or Locality.

2. Perform a *single level* search for Application Entities which support the required application contexts. This assumes that all services which are supporting default access for the organisation are registred at one level below (possibly by the use of aliaes), and that other services (specific machines or parts of the organisation) are represented further down the tree. This seems to be a reasonable layout, and its utility can be evaluated by experiment.

8 Experience

An experimental implementation of this has been written by Colin Robbins. The example in Figure 2 shows that it can be very effective at locating known individuals with a minimum of effort.

9 Relationship to other work

Colin Robbin's work on the interface "Tom" and implementation of a distribution list interface strongly influenced this proposal.

Some of the ideas used here originally came from a UK Proposal to the ISO/CCITT Directory Group on "New Name Forms" [Kil89]. This defined, and showed how to implement, four different types of names:

Typed and Ordered The current Distinguished Name is a restricted example of this type of name.

will be made, and the user prompted for "Joseph Z. Bloggs" as the only possible match.

If "Computer Science" is supplied, first

```
Physics, University College London, GB
```

will be searched, and the user will reject the approximate match of "Colin Skin". Then a subtree search of

```
University College London, GB
```

will be made, looking for a person. Then a single level search will be made looking for Org Unit, and

```
Computer Science, University College London, GB
```

will be returned without prompting (exact match).

Supplying "Steve Kille" will lead to a failed subtree search of

```
Physics, University College London, GB
```

and lead straight to a subtree search of

```
University College London, GB
```

This will lead to an exact value match, and so a single entry returned without prompting.

If "Andrew Findlay, Brunel" is supplied, the first element of the environment will be skipped, single level search of "Brunel" under "GB' will find:

```
Brunel University, GB
```

and a subtree search for "A Findlay" initiated. This will yield

```
Andrew Findlay, Computing and Media Services, Brunel University, GB

Dr A J Findlay, Manufacturing and Engineering Systems,
Brunel University, GB
```

and the user will be prompted with a choice.

This approach shows how a simple format of this nature will "do the right thing" in many cases.

6 Support required from the standard

Fortunately, all that is needed is there! It would be useful to have "friendly country name" as a standard attribute.

These allow some optimisation, based on legal choices of schema. Keeping filters short is usually desirable to improve performance.

A few examples of this, where a base object has been determined (either by being the environment or by partial resolution of a purported name), and the next element of a purported name is being considered. This will generate a single level search. What varies is the types being filtered against. If the DN is:

```
University College London, GB
```

The search should be for Org Unit or Locality. If the DN is:

```
Organisation=UN
```

the search should be for Org Unit or Locality.

There may be some improvements with respect to very short keys. Not making approximate or substring matches in these cases seems sensible[1].

4.5 Bottom level

The "Bottom Level" is to deal with leaf entries in the DIT. This will often be a person, but may also be a role, an application entity or something else.

The last component of a purported name may either reference a leaf or non-leaf. For this reason, *both* should be tested for. As a heuristic, if the base object for the search has two or more components it should be tested first as a bottom level name and then intermediate. Reverse this for shorter names. This optimises for the (normal) case of non-leaves high up the tree and leaves low down the tree.

For bottom level names, make an approximate and substring match against Common Name, Surname, and User ID. Where common name is looked for, A full subtree search will be used when at the second level of the DIT or lower, otherwise a single level search.

For example, if I have resolved a purported name to the distinguished name

```
University College London, GB
```

and have a single component `Bloggs`, this will generate a subtree search.

5 Examples

This is all somewhat confusing, and a few examples are given. These are all in the context of the environment shown in Table 2 on Page 7.

If "Joe Bloggs" is supplied, a subtree search of

```
Physics, University College London, GB
```

[1] It might be desirable to allow "*" as a part of the purported name notation

4. Automatically move to the next element of the environment if no matches are found.

When the final component is matched, a set of names will be identified. If none are identified, proceed to the next environment element. If the user rejects all of the names, processing of the next environment element should be confirmed.

The exact approach to matching will depend on the level of the tree at which matching is being done. We can now consider how attributes are matched at various levels of the DIT.

There is an issue of approximate matching. Sometimes it helps, and sometimes just returns many spurious matches. When a search is requested, all relevant attributes should be returned, so that distinguished and non-distinguished values can be looked at. This will allow a distinction to be made between good and poor matches. It is important that where, for example, an acronym exactly matches an organisation, that the user is not prompted about other organisations where it matches as a substring.

4.3 Top Level

In this case, a match is being done at the root of the DIT. Three approaches are suggested, dependent on the length of supplied name. All lead to a single level search of the top level of the DIT.

Exactly 2 This is assumed to be a 3166 two letter country code, or an exact match on a friendly country or organisation (e.g., UK or UN). Do exact match on country and friendly country.

Greater than 2 Make an approximate and substring match on friendly country and organisation.

4.4 Intermediate Level

Once the root level has been dealt with, intermediate levels will be looking for organisational components (Organisation, Locality, Org Unit). In some cases, private schema control will allow the system to determine which is at the next level. In general this will not be possible. In each case, make a substring and approximate match search of one level. The choice depends on the base object used in the search.

1. If DN has no Organisation or Locality, filter on Organisation and Locality.
2. If DN has Org Unit, filter on Org Unit.
3. If DN has Organisation, filter on Locality and Org Unit.
4. If DN has Locality, filter on Organisation.

4.2 Matching

A purported name will be supplied, usually with a small number of components. This will be matched in the context of an environment.

Where there are multiple components to be matched, these should be matched sequentially. If an unambiguous DN is determined, the match continues as if the full DN had been supplied. For example if

```
Stephen Kille, UCL
```

is being matched in the context of environment `GB`, first `UCL` is resolved to the distinguished name:

```
University College London, GB
```

Then the next component of the purported name is taken to determine the final name. If there is an ambiguity (e.g., if `UCL` had made two matches, both paths are explored to see if the ambiguity can be resolved. Eventually a set of names will be passed back to the user.

Each component of the environment is taken in turn. If the purported name has more components than the maximum depth, the environment element is skipped. The advantage of this will be seen in the example given later.

A match of a name is considered to have three levels:

Exact A DN is specified exactly

Good Initially, a match should be considered good if it is unambiguous, and exactly matches an attribute value in the entry. For human names, a looser metric is probably desirable (e.g., `S Kille` should be a good match of `S. Kille`, `S.E. Kille` or `Steve Kille` even if these are not explicit alternate values).

Poor Any other substring or approximate match

Following a match, the reference can be followed, or the user prompted. If there are multiple matches, more than one path may be followed. There is also a shift/reduce type of choice: should any partial matches be followed or should the next element of the environment be tried. The following heuristics are suggested, which may be modified in the light of experience. The overall aim is to resolve cleanly specified names with a minimum of fuss, but give sufficient user control to prevent undue searching and delay.

1. Always follow an exact match.
2. Follow all good matches if there are no exact matches.
3. If there are only poor matches, prompt the user. If the user accepts one or more match, they can be considered as good. If all are rejected, this can be treated as no matches.

Number of Components	Environment
1	Physics, University College London, GB University College London, GB GB –
2	GB University College London, GB –
3+	– GB University College London, GB

Table 2: Local environment for private DUA

Number of Components	Environment
1,2	US CS –
3+	– US CA

Table 3: Local environment for US Public DUA

In most cases, the environment will start at a local point in the DIT, and move upwards. Examples are given in Tables 2 and 3. Table 2 shows an example for a typical local DUA, which has the following characteristics:

One component Assumed first to be a user in the department, then a user or department within the university, the a national organisation, and finally a country.

Two components Most significant component is first assumed to be a national organisation, then a department (this might be reversed in some organisations), and finally a country.

Three or more components The most significant component is first assumed to be a country, then a national organisation, and finally a department.

```
Steve Keill, Comp Sci, Univarstiy College London, GB
```

Friendly Country A "friendly country name" can be used instead of the ISO 3166 two letter code. For example: UK; USA; France; Deutchland.

Type Changing A type may be omitted, even where it does not follow the hierarchy. Type variants can be explored. The previous distinguished name:

```
James Hacker
locality=Basingstoke
Widget Inc
GB
```

Could be represented by the purported name:

```
James Hacker
Basingstoke
Widget Inc
GB
```

4 Matching a purported name

The major utility of the purported name is to provide the important "user friendly" characteristic of guessability. A user will supply a purported name to a user interface, and this will be resolved onto a distinguished name. When a user supplies a purported name there is a need to derive the DN. In most cases, it should be possible to derive a single name from the purported name. In some cases, ambiguities will arise and the user will be prompted to select from a multiple matches. This should also be the case where a component of the name did not "match very well".

There is an assumption that the user will simply enter the name correctly. The purported name variants are designed to make this happen! There is no need for fancy window based interfaces or form filling for many applications of the directory. Note that the fancy interfaces still have a role for browsing, and for more complex matching. This type of naming is to deal with cases where information on a known user is desired and keyed on the user's name.

4.1 Environment

All matches occur in the context of a local environment. The local environment defines a sequence of name of a non-leaf objects in the DIT. This environment effectively defines a list of acceptable name abbreviations where the DUA is employed. It also defines an order in which to operate.

This list is defined in the context of the number of name components supplied. This allows varying heuristics, depending on the environment, to make the approach have the "right" behaviour.

```
Common Name, (((Organisational Unit)*,  Organisation,) Country)
```

Explicitly typed RDNs may be inserted into this hierarchy. Two types of object are named. They are distinguished by context in the application:

Common Name Object Typically a leaf object and a person. In this case, the LHS RDN is assumed to be common name, and other RDNs follow the hierarchy.

Other Object Typically a non-leaf object.

3 Purported names

A purported name is what a user supplies to an interface for resolution into one or more distinguished names. The same notation is used as for distinguished name. A purported name may differ from a distinguished name in a number of ways which are itemised below.

Typically, an interface will display a distinguished name, using the distinguished name notation. However, it may display a purported name, as this will be more pleasing to the user. The most common example of this is where the higher components of the distinguished name are not displayed (abbreviation).

The ways in which a purported name may vary from a distinguished name are now described:

Abbreviation Some of the more significant components of the DN will be omitted, and then defaulted in some way (e.g., relative to a local context). For example:

```
Steve Kille
```

Component Omission An intermediate component of the name may be omitted. Typically this will be an organisational unit. For example:

```
Steve Kille, University College London, GB
```

In some cases, this can be combined with abbreviation. For example:

```
Steve Kille, University College London
```

Approximation Approximate renditions or alternate values of one or more of the components will be supplied. For example:

```
Stephen Kille, CS, UCL, GB
```

or

```
<name> ::= <name-component>
       | <name-component> <spaced-separator> <name>

<spaced-separator> ::= <optional-space>
                       <separator>
                       <optional-space>

<separator> ::= "," ( <CR> )
              | <CR>

<optional-space> ::= *( " " )

<name-component> ::= <attribute>
        | <attribute> <optional-space> "+" ( <CR> )
          <optional-space> <name-component>

<attribute> ::= <string>
        | <key> <optional-space> "=" <optional-space> <string>

<key> ::= 1*( <keychar> )
<keychar> ::= letters, numbers, and space

<string> ::= *( <stringchar> | <pair> )
         | '"' *( <stringchar> | "," | "+" | "=" | <CR> ) '"'

<special> ::= "," | "=" | '"' | <CR> | "\" | "+"
<pair> ::= "\" <special>
<stringchar> ::= any char except <special>
```

Figure 1: BNF Grammar for Distinguished and Purported Name

Key	Attribute (X.520 keys)
Locality	LocalityName
Area	LocalityName
Organisation	OrganizationName
Organization	OrganizationName

Table 1: Standardised Keywords

```
Steve Kille
Computer Science
University College London, GB
```

Another name might be:

```
Christian Huitema, INRIA, FR
```

An example, showing how different attribute types are handled:

```
James Hacker
locality=Basingstoke
Widget Inc
GB
```

An example showing quoting of a comma in an Organisation name:

```
L. Eagle, "Sue, Grabbit and Runn", GB
```

2.3 Formal definition

The structure is specified in a BNF grammar in Figure 1.

The quoting mechanism is used for the following cases:

- Strings containing ",", "+", "="or "" or `<CR>`
- Strings with leading or trailing spaces
- Strings containing consecutive spaces

A list of valid keywords for well known attribute types used in naming is given in Table 1. This is a list of keywords which must be supported. These are chosen because they appear in common forms of name, and can do so in a place which does not correspond to the default schema used. If other attributes are used for naming, this can always be extended locally.

Only string type attributes are considered, but other attribute syntaxes could be supported locally. It is assumed that the interface will translate from the supplied string into PrintableString or T.61.

The "+" notation is used to specify multi-component RDNs. In this case, the types for attributes in the RDN must be explicit.

The name is presented/input in a little-endian order. If no type are given, a type hierarchy (schema) is assumed as:

1 Why a notation is needed

Many OSI Applications make use of Distinguished Names (DN) as defined in the OSI Directory [CCI88]. The main reason for having a notation for is for interacting with a user interface. This proposal is coming dangerously close to the sin of standardising interfaces. However, there are aspects of presentation which it is desirable to standardise.

In very many cases, a user will be required to input a name. This notation is designed to allow this to happen in a uniform manner across many user interfaces. The intention is that the name can just be typed in. There should not be any need to engage in form filling or complex dialogue.

A secondary goal, is to have a common format to be able to unambiguously refer to names. This notation should *not* be needed to write on business cards or in the minutes of meetings. If this is the case, the directory has failed to a large extent. It should be possible to take the "human" description given at the meeting, and use it directly. The means in which this happens will become clear later. In practice, it will often be useful to explicitly write down directory names, even though it is not strictly needed.

2 A notation for Distinguished Name

2.1 Goals

The following goals are laid out:

- It should be intuitive for the majority of names encountered
- It should be fully general
- It should be possible to lay it out in a number of ways

2.2 Informal definition

This notation is designed to be convenient for common forms of name. Some examples are given. My directory (distinguished) name would be written:

```
Steve Kille, Computer Science, University College London, GB
```

This may be folded, perhaps to display in multi-column format. For example:

```
Steve Kille,
Computer Science,
University College London,
GB
```

or

Contents

1 Why a notation is needed **2**

2 A notation for Distinguished Name **2**
2.1 Goals 2
2.2 Informal definition 2
2.3 Formal definition 3

3 Purported names **5**

4 Matching a purported name **6**
4.1 Environment 6
4.2 Matching 8
4.3 Top Level 9
4.4 Intermediate Level 9
4.5 Bottom level 10

5 Examples **10**

6 Support required from the standard **11**

7 Support of OSI Services **12**

8 Experience **12**

9 Relationship to other work **12**

10 Proposal **14**

A Pseudo-code for the matching algorithm **16**

List of Figures

1 BNF Grammar for Distinguished and Purported Name 4
2 Example usage of User Friendly Naming 13
3 Matching Algorithm 19

List of Tables

1 Standardised Keywords 3
2 Local environment for private DUA 7
3 Local environment for US Public DUA 7

Research Note RN/90/29
June 1990

Research
Note

Using the OSI Directory to achieve User Friendly Naming

S.E. Kille

Abstract

The OSI Directory has user friendly naming as a goal. A simple minded usage of the directory does not achieve this. Two aspects not achieved are:

- A user oriented notation
- Guessability

This proposal sets out some conventions for representing names in a friendly manner, and shows how this can be used to achieve really friendly naming.

Department of Computer Science
University College London