# Hashtable

## Lösungsidee

### Hashtable

Bei der Realisierung der *Hashtable* sollte die vorgegeben Header Datei verwendet werden. Dabei werden alle grundlegenden Methoden implementiert, welche für eine *Hashtable* nötig sind. Die zu implementierenden Funktion sind in der Angabe beschrieben.Alle Elemente die neu hinzugefügt werden innerhalb der Hashtable in folgende Datenstruktur abgespeichert:

* *vector<list<value\_type>>*

Der *value\_type* ist ein Template Parameter der im beim instanziieren der Klasse gewählt wird. Der Vorteil bei dieser Datenstruktur ist, dass der *vector* die Eigenschaft hat, dass es leicht ist ihn zu vergrößern bzw. zu verkleinern und es die Möglichkeit gibt, auf Elemente innerhalb dieses Vektors mittels Index zuzugreifen. Die Elemente dieses Vektors wurden als *list* implementiert, da diese im Prinzip nur durchlaufen werden und nach dem gewünschten Element durchsucht werden. Um einen möglichst schnellen Zugriff auf die einzelnen Elemente zu ermöglichen, wird mittels einer im Konstruktor übergebenen Hash-Funktion ein Hash generiert. Der Einfachheit halber wird bei dieser Hashtable für die Index des aktuellen Elementes der Wert des Elementes der Hash Funktion übergeben, dieser Hash dann modulo der Größe des Vektors gerechnet, dass ein vorhandener Index gefunden wird. Diese Logik für das Berechnen dieses Hash Wertes sollte in eine eigene Funktion namens *getHashedValue*() ausgelagert werden.Beim Hinzufügen eines Elementes wird dazu vorher überprüft ob das Element bereits in der Hashtable vorhanden ist, wenn nein wird der Hash Wert wie oben beschrieben berechnet und danach mit dem zurückgegebenen Index die Liste im Vektor selektiert welcher das Element angefügt werden sollte. Dieser Liste wird das Element hinten angefügt und anschließend die aktuelle Größe erhöht. Wenn diese Größe einen gewissen Wert übersteigt, in unserem Fall sollte das der *MaxLoadFactor* mal der Kapazität sein, wird die Funktion *rehash* mit der verdoppelten aktuellen Größe aufgerufen. In der Funktion *rehash* sollte die Tabelle vergrößert und anschließend neu befüllt werden. Dazu wird zu Erst eine Kopie der aktuellen Tabelle angelegt, anschließend wird die Kapazität mit der gewünschten Größe überschrieben und die aktuelle Größe auf 0 gesetzt. Die Tabelle wird anschließend auf die neue Kapazität geändert. Dies sollte mit Hilfe der *resize* Funktion geschehen. Diese Tabelle wird anschließend geleert und dann neu befüllt. Aufgrund der Tatsache, dass sich die Größe der Tabelle geändert hat, wird jetzt auch ein neuer Hash Wert für die einzelnen Elemente generiert, daher muss die Tabelle neu befüllt werden.

Die erase Funktion sollte gleich aufgebaut sein wie die insert Funktion. Statt der Überprüfung ob das Element nicht vorhanden ist, sollte überprüft werden, ob das Element vorhanden ist, wenn nicht kann es auch nicht entfernt werden und es muss daher nichts passieren. Falls ein Element mit dem Wert vorhanden ist, wird die Liste mit dem Element mit dem Wert den wir aus der Funktion *getHashedValue* bekommen (wie oben beschrieben) selektiert. Da wir uns sicher sein können, dass der Wert in dieser Liste existiert, brauchen wir anschließend nur noch die *remove* Funktion auf die selektierte Liste aufrufen. Danach wird die aktuelle Größe wieder um eines verringert. Wenn diese Größe unter einen bestimmten Wert sinkt, muss wieder die *rehash* Funktion aufgerufen werden. Der bestimmte Wert ergibt sich dabei aus der Multiplikation von *MinLoadFactor* und der aktuellen Kapazität. Die Funktion *rehash* wird dabei mit der halbierten aktuellen Größe aufgerufen und funktioniert wie oben beschrieben.

Die *contains* Funktion sollte feststellen, ob ein bestimmtes Element in der Tabelle vorhanden ist. Dazu wird über die *getHashedValue* der Hash Wert für das aktuelle Element generiert, über diesen Wert wird schließlich die Liste aus dem Vektor selektiert und diese Liste wird danach durchsucht ob der Wert enthalten ist. Wenn das Ergebnis dieser Suche dem End-Iterator der Liste entspricht existiert es nicht ansonsten gibt es ein Element.

Bei der Überladung des *== Operators* sollte zu Erst überprüft werden ob die übergebene Tabelle die gleiche Adresse hat, wie die aktuelle Klasse. Dann ist es in jedem Fall das gleiche Objekt bzw. hat die gleichen Elemente. Wenn die Größe der übergebenen Tabelle sich von der aktuellen Tabelle unterscheidet kann in jedem Fall gesagt werden, dass die Tabellen unterschiedlich sind. Falls diese beiden Fälle nicht eintreten, sollte die übergebene Liste in einer Schleife durchlaufen werden und für jedes Element dieser Tabelle überprüft werden, ob es im aktuellen Objekt vorhanden ist. Falls nicht kann davon ausgegangen werden, dass es nicht vorhanden ist. Falls die Schleife nicht abgebrochen wurde, sind alle Elemente der übergebenen Tabelle vorhanden und es kann daher davon ausgegangen werden, dass diese Tabellen zumindest die gleichen Elemente besitzen. Für diese Funktion wird es nötig sein einen Iterator zu implementieren. Wie dieser aussehen sollte wird weiter unten beschrieben.

Die Überladung des *<< Operators* geht in einer Schleife alle Elemente der übergebenen Tabelle durch und gibt diese auf dem übergebenen *ostream* aus. Auch diese Funktion benötigt die Implementierung des Iterators.

Die weiteren Funktionen sind trivial und sollten entweder einzelne Werte zurückgeben oder eine Kombination dieser.

### Iterator

Bei der Implementierung des Iterators für die Hashtable sollte darauf geachtet werden, dass alle Elemente nacheinander durchlaufen werden. Das heißt, wir müssen nacheinander alle Listen aus dem Vektor selektieren und die einzelnen darin enthaltenen Elemente selektieren. Beim instanziieren des Iterators sollte hierzu auf das erste gültige Element in der Hashtable verwiesen werden. Das erste gültige Element wird dabei dadurch gefunden, dass man mittels vector.begin() startet und solange die zurückgegebene Liste leer ist den dabei zurückgegeben Iterator erhöht, bis die erste Liste kommt welche mindestens ein Element enthält. Sobald dieses Gefunden wird, werden zwei Membervariablen in der Iterator Klasse gesetzt. Diese Funktionalität sollte in der *begin* Funktion realisiert werden. Falls das Ende der Hashtable gewünscht ist, wird der gleiche Prozess durchgeführt, nur wird vom Ende des Vektors ausgegangen und so lange verringert bis die aktuelle Liste mindestens ein Element enthält. Die Membervariablen werden dann auf diesen Iterator mit der ersten Liste gesetzt und der zweite Iterator wird auf das letzte Element dieser Liste gesetzt. Diese Funktionalität sollte in einer Funktion end() realisiert werden.

Falls auf den Iterator der Hashtable *operator++* augerufen wird, wird zu Erst der Listeniterator erhöht. Wenn dieser am Ende der Liste angekommen ist, dann wird der Vektoriterator erhöht und der Listeniterator danach auf den Anfang der aktuellen Liste gesetzt. Dieser Vorgang wird so lange durchgeführt bis eine Liste gefunden wurde die mindestens ein Element hat.

Wenn *operator--* aufgerufen wird, sollte zu Erst überprüft ob der *listIterator* dem Beginn des Vektoriterators entspricht. Wenn nicht sollte er verringert werden. Falls diese Aussage wahr ist, wird der *vectorIterator* solange verringert, bis er entweder dem Anfang des Vektors entspricht oder die Liste auf die der Vektoriterator verweist eine Größe größer als 0 hat. Falls der Vektoriterator also dem Beginn des Vektors entspricht sollte die oben beschrieben Funktion *begin* aufgerufen werden. Falls nicht, wird der *listIterator* an das Ende des Vektoriterators gesetzt und um eins verringert.

Bei der Überladung des == Operators sollte überprüft werden ob der *vectorIterator* dem übergeben *vectorIterator* entspricht und ob der *listIterator* dem übergebenen entspricht.

Bei dem != Operator sollte überprüft werden ob der *vectorIterator* ungleich dem übergeben *vectorIterator* ist oder der *listIterator* ungleich dem übergebenen entspricht.

Die Funktionen für die *reference* und den *pointer* sind trivial und benötigen daher keine nähere Beschreibung.

Es sollten auch die *begin*() *end*() bzw. die *cbegin*() und *cend*() Funktionen für die Hashtable implementiert werden. Dabei sollte für die *begin* Funktion einfach eine Instanz des zu implementierenden *const\_iterators* erstellt weden und bei der *end* Funktion ebenfalls. Der Unterschied ist, dass bei der end Funktion der Iterator noch an das Ende gesetzt wird. Die Funktionen *cbegin* und *cend* sowie deren Überladungen rufen im Prinzip nur *begin* und end auf.